DOCUMENT RESUME

ED 036 900 EA 002 792

AUTHOR	Hausdorff, Henry, Ed.
TITLE	American Educational Research Association Paper Abstracts, 1968 Annual Meeting.
INSTITUTION	American Educational Research Association, Washington, D.C.
PUB DATE	Feb 68
NOTE	374p.; Abstracts of papers presented at the Paper Sessions of the Annual Meeting of the AERA (Chicago, Illinois, Feb. 8-10, 1968)
EDRS PRICE	EDRS Price MF-$1.50 HC-$18.80
DESCRIPTORS	*Abstracts, Classroom Communication, Concept Teaching, *Conference Reports, Conservation Education, Culturally Disadvantaged, *Educational Research, Educational Strategies, Evaluation, Individual Differences, *National Organizations, Programed Instruction, Reinforcement, Socioeconomic Status, Student Characteristics, Teacher Behavior, Testing Problems, Verbal Communication

ABSTRACT

This document contains abstracts of approximately 400 papers presented at the 1968 annual meeting of the American Educational Research Association. A sampling of the varied subjects covered includes: classroom interaction, verbal behavior, concept learning, conservation concepts, teaching the disadvantaged, teaching strategies, research on testing, characteristics of college populations, vocational education, career development, planning and decision making in administrative organization, and research on classroom behavior. A topical index of the papers and an index of participants conclude the collection. Related documents are EA 002 791, EA 002 793, and EA 002 794. (DE)

ED036900

N-Y

A.E.R.A. PAPER ABSTRACTS

HENRY HAUSDORFF
Editor

Paper Sessions of the

1968 ANNUAL MEETING

EA 002 792

American Educational Research Association
1126 Sixteenth Street, N.W.
Washington, D.C. 20036

$2.00
($2.25 by
mail)

ED036900

AMERICAN EDUCATIONAL RESEARCH ASSOCIATION

PAPER ABSTRACTS

Editor
HENRY HAUSDORFF

Paper Sessions of the

1968 ANNUAL MEETING

CHICAGO
CONRAD HILTON HOTEL
FEBRUARY 8-10

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

American Educational Research Association
1126 Sixteenth Street, N.W.
Washington, D.C. 20036

$2.00
($2.25 by mail)

Officers of the American Educational Research Association Concerned with the 1968 Program

PRESIDENT	JOHN I. GOODLAD, University of California at Los Angeles
VICE-PRESIDENTS FOR DIVISIONS	
A	JACK A. CULBERTSON, Ohio State University
B	B. OTHANEL SMITH, University of Illinois
C	ROBERT GLASER, University of Pittsburgh
D	JACK C. MERWIN, University of Minnesota
E	JOHN D. KRUMBOLTZ, Stanford University
EXECUTIVE OFFICER	RICHARD A. DERSHIMER, 1126 Sixteenth Street, N.W. Washington, D.C. 20036
PROGRAM COMMITTEE	
Program Chairman	ROBERT GLASER, University of Pittsburgh
Vice-Chairman	PAUL M. KJELDERGAARD University of Pittsburgh
Division A	SAMUEL GOLDMAN, Syracuse University
Division B	ARNO B. BELLACK, Columbia University
Division C	RICHARD C. ANDERSON, University of Illinois
Division D	JOHN E. MILHOLLAND, University of Michigan
Division E	T. ANTOINETTE RYAN, Oregon State University
N.C.M.E.	J. THOMAS HASTINGS, University of Illinois
Graduate Student Seminar	RICHARD L. TURNER, University of Indiana
Pre-Sessions	GENE V. GLASS, University of Colorado

A.E.R.A. Division Program Committees

1967-1968

The following A.E.R.A. members served as evaluation teams for the contributed papers abstracted in this volume:

A.E.R.A. PROGRAM CHAIRMAN—ROBERT GLASER

DIVISION A SAMUEL GOLDMAN, Chairman
DALE L. BOLTON
EDWIN M. BRIDGES
WILLIAM G. MONAHAN

B ARNO A. BELLACK, Chairman
PAUL KOMISAR
KENNETH J. REHAGE
MARY BUDD ROWE
LOUIS M. SMITH

C RICHARD C. ANDERSON, Chairman
WELLS HIVELY II
JOANNA P. WILLIAMS
JOSEPH L. WOLFF
LAWRENCE M. STOLUROW
WALTER H. MacGINITIE
FREDERICK J. McDONALD

D JOHN L. MILHOLLAND, Chairman
JOAN BOLLENBACHER
WILLARD G. WARRINGTON

E T. ANTOINETTE RYAN, Chairman
ISADORE GOLDBERG
RAYMOND C. HUMMEL
THOMAS P. NALLY

TABLE OF CONTENTS

Page
OFFICERS OF A.E.R.A. 2
A.E.R.A. DIVISION OF PROGRAM COMMITTEES 3
TABLE OF CONTENTS 4
EDITOR'S COMMENT 6

PAPER ABSTRACTS

SESSIONS

1.1 Computer-Related Research Models 8
1.6 Classroom Interaction I 11
1.7 Attitudes Toward Education 15
1.8 Verbal Behavior I 19
1.10 Planning and Decision-Making in Educational Organization 22
1.11 Political Aspects of Educational Policy Determination 27
2.2 Predicting Academic Success 31
2.5 Learning and Individual Differences I 35
2.6 Contributed Papers I 38
3.4 Learning from Prose I 42
3.5 Early Childhood Education 46
3.6 Science and Mathematics 49
3.7 Social Studies 53
4.4 Research on Classroom Behavior 57
4.6 Special Problems in Administrative Research 61
4.7 Supervisory Behavior and Teacher Motivation 64
4.8 Concept Learning I 68
5.3 Beginning Reading 72
5.8 Verbal Behavior II 76
5.9 Conservation Concepts I 80
6.2 Teaching the Disadvantaged I 84
6.4 Programed Instruction I 88
6.5 Research on Special Groups 92
6.6 Behavioral Counseling 96
6.7 Self-Concept and Pupil Performance 100
7.2 Microteaching 104
7.3 Research on Testing I 108
7.4 Development of Educational Measuring Instruments 112
7.5 Socioeconomic Status I 116
7.6 Reading Research 120
7.7 Language Arts 124
7.8 Compensatory and Vocational Education 128
7.9 Group Characteristics 132
8.3 Mathematics Instruction 135
8.6 Programed Instruction II 139
8.11 Contributed Papers II 143
8.12 Characteristics and Behaviors of School Principals 147
9.1 Teacher Behavior 152
9.2 Reinforcement I 157
9.3 Memory 162
10.5 Evaluation of Teaching Strategies I 166
10.8 Counseling: Evaluation and Micro-Procedures 169
10.9 Verbal Behavior III 173

Page
10.10 Dimensions of Individual Difference 176
10.11 Contributed Papers 180
11.1 Learning from Prose II 184
11.3 Organizational Theory and Administrative Behavior 188
11.4 Research on Testing II 192
11.5 Concept Learning II 196
11.7 Contributed Papers I 200
12.5 Language Instruction for Disadvantaged Students 204
12.6 Psycho-Social Development 208
12.7 Socioeconomic Status II 212
12.8 Teacher Behavior and Organizational Structure 216
13.2 Curricula for Disadvantaged Students 220
13.3 Statistical Techniques 225
13.4 Measurement of Complex Intellectual Processes 229
13.5 Verbal Behavior IV 233
13.6 Learning and Individual Differences II 236
14.1 Techniques for Interpretation of Results 241
14.5 Concept Learning III 244
14.6 Programed Instruction III 248
14.8 Verbal Behavior V 252
14.9 Research on Testing III 256
14.10 Contributed Papers II 260
15.1 Problem Solving and Thinking 264
15.2 Reinforcement II 268
15.3 Research on Teaching 272
15.4 Experimental Design 276
16.3 School Board-Superintendent Relationship 279
16.4 Test Theory 283
16.6 Growth Relationships and Pupil Development 287
17.4 Teaching the Disadvantaged II 291
17.6 Variables Related to Educational Outcomes 295
17.7 College Students: Values and Attitudes 298
18.6 Evaluation of Teaching Strategies II 301
18.7 Curriculum Development 304
18.8 Conservation Concepts II 308
18.9 Characteristics of College Populations 312
19.3 Verbal Behavior VI 316
19.4 Studies of Prospective Teachers 319
19.5 Empirical Test Research 322
19.6 Vocational Interests and Career Development 326
20.1 School Integration 329
20.2 Classroom Interaction II 333
20.3 Studies of Learning Environments 337
20.4 Curriculum Evaluation 341
20.5 Characteristics of Students 347
20.6 College Student Adjustment and Development 350
20.7 Studies of Non-American Subjects 354
20.8 Teacher Characteristics 358
20.9 Concept Learning IV 362
TOPICAL INDEX OF PAPER SESSIONS 366
INDEX OF PARTICIPANTS 368

EDITOR'S COMMENT

The 1968 edition of *"A.E.R.A. Paper Abstracts"* represents the most ambitious undertaking of this publication, with close to 400 papers summarized. Of course, symposia, invited addresses, and special events, which comprise a substantial portion of the 1968 program, have not been abstracted.

The splendid format developed by Ellis B. Page for the 1967 edition has been essentially retained, as it is a highly commendable model. One change has been initiated in the indices at the back of the volume. Instead of utilizing page numbers for reference, session numbers are furnished. The abstracts are presented sequentially by session, and each page of the text has the session number in the upper right-hand corner. This new referral device, conforming with the chronological order of sessions, should add to the ease and convenience of identifying author and abstract.

It should be noted that for the most part, the abstracts appear as presented. Regretfully, however, a substantial number required considerable editing to have them comply with length requirements. In addition, quite a few were submitted in well-nigh illegible carbon copies or inferior photoreproductions. This made the editor's task more difficult, and may be responsible for unintentional injustices to the author's work.

It would be inexcusable if this opportunity were not taken to acknowledge the particular assistance of a few people. Without their help, this endeavor could not have been completed on time. Gary E. Hanna of the A.E.R.A. Washington office greatly facilitated communications and adherence to deadlines. Mrs. Pamela H. Greene assisted with essential typing, and received wholehearted cooperation from all staff members on other details. Untold indebtedness is due to David M. Hausdorff for his tireless efforts and assistance in all aspects of the editorial work including editing, proofreading, and typing. Appreciation is acknowledged to Robert Glaser, 1968 A.E.R.A. Program Chairman, for having extended the privilege of participation in this program.

HENRY HAUSDORFF
University of Pittsburgh

January, 1968

The A.E.R.A. Paper Abstracts for 1968

The following abstracts are arranged in order of presentation during the Annual Meeting. They include only those contributed papers submitted to the Divisional Program Chairman, and accepted by the evaluation teams. In most cases they are as written by the authors, but in some abstracts changes have been made to shorten the prose to an acceptable length, or to make greater consistency in reporting practices.

SPECIAL NOTES: The person reading the paper will usually be the first author listed. Where someone else was scheduled to read it, he will be denoted by an asterisk (*).

At times papers are written and presented by researchers who are not Members of A.E.R.A. In such cases, the name of the Member sponsor will appear, in parentheses, after that of the author.

Session numbers are not in a complete series. They are rather used to designate certain time blocks, and many numbers are not used at all.

9:00 to 9:55 A.M. **SESSION 1.1** **DIVISION D**

COMPUTER-RELATED RESEARCH MODELS

CHAIRMAN:

Ellis B. Page, University of Connecticut

PAPERS:

Some Additional Developments in a Computer Model of Concept Attainment

FRANK B. BAKER, University of Wisconsin

For some time, the author has been developing a computer model of the concept attainment process and earlier had suggested that one might be able to develop computer program modules from which given simulations could be constructed. However, at that time it was not clear as to what these modules would be or how one would utilize them. In recent attempts to study the process by which a human subject translates the experimenter's initial instructions into overt and covert behavior, a possible approach to these modules becomes apparent. The simulation computer program has been restructured to include a hierarchy of "contexters" whose task was to construct a learning strategy and fractionate this strategy into executable behavior. Due to the initial design of the computer model and its subsequent restructuring, it was quite easy to describe the operational aspects of the program in simple English. The simplicity was further enhanced by a deliberate effort to minimize the number of verbs, nouns, etc. involved in the descriptions. Then a one-to-one correspondence between the verbs and computer subroutines was established. Thus, the basic operational module became a verb. However, verbs must have operands and both the verbs and operands can have modifiers. Given the verb as the basic module, the next upper level of "contexting" was devoted to parsing an English language description of the behavior and providing the verb with its operands and modifiers. Under such a system one need only provide the computer program with a simple English sentence and the sentence can be translated into the appropriate specific behavior. Converting several sentences could result in a meaningful sequence of behaviors. The remainder of the present paper was devoted to illustrating the operational aspects of the concept attainment process using the linguistic approach.

◆ ◆ ◆

A Simplified Program for N-Way Analysis of Variance

TOM HOUSTON, University of Wisconsin

The purpose of this paper is to describe a computer program which the author has developed to facilitate the correct use of the analysis of variance (ANOVA). The program is intended to minimize effort and the likelihood of error on the part of users of this statistical technique.

In his 1966, *Psychometrika* article "Experimental design symbolization and model derivation," Wayne Lee proposed a simple notational system for representing factorial designs, and provided an analytic procedure to derive the statistical model from the symbolic formula. The present paper describes an alternative and computationally simpler procedure which obtains sources of variation from Lee's algebraic formula for any balanced, fully replicated or nested factorial design. In essence, all 2^n sources of variation which are possible in an n-factor design are generated; those which because of nesting cannot be separately estimated are subsequently discarded. Except in a minority of designs where most of the independent variables appear as nested factors, this "subtractive" process is more efficient than Lee's "additive" approach.

This derivation procedure, in conjunction with an extremely rapid numerical scheme, has been cast in the form of a computer program which provides n-way ANOVA tables for a large class of factorial designs. To perform an ANOVA, the researcher must supply the program with his data, a Lee-type design formula, and only one additional control card.

As an example of a design formula, the design given in section 12.20 of Hayes, *Statistics for Psychologists,* pp. 385, 6, can be represented as:

REPLICATIONS, 10 (NORMS, 2 * STANDING, 3).

The numerals give the number of levels of each factor, while the asterisk and parentheses respectively indicate crossed and nested relationships among factors.

◆ ◆ ◆

Identification of Homogeneous Groups of Students by Computer

M. CLEMENS JOHNSON, The University of Michigan

This paper describes the design and discusses a research application of a computer program for identifying homogeneous groups or clusters of students. The research methodology represented in the program is new to educational stdudies, providing an alternative to the usual approach of comparing previously contrived groups. Output data include clusters of individuals formed by the computer in successive iterations through a sample. Measurements may consist of test scores, item scores, rating scale responses, and similar variables.

Individuals within a cluster tend to be alike based upon computed similarity coefficients which in turn are based upon an item by item comparison of available numerical data. However, additional study is needed to identify characteristics which differentiate the various clusters. The clusters should be viewed chiefly as exploratory, providing a starting point for further investigation.

In general the program can be valuable in attaining some compatibility between the desire to theorize about groups of individuals and the need to explain widespread individual differences. For instance, instead of comparing previously defined groups, such as boys and girls, high and low achievers, or white and Negro students, the program can be used to identify subgroups, mixed or pure, which are more alike than individuals at random from either or both groups.

◆ ◆ ◆

A Computerized Determination of the Readability of Programmed Materials Using Complete Units

ROBERT A. SHAW and MILTON D. JACOBSON, University of Virginia

The purposes of this study were: (1) to automate a procedure which could be used to obtain total variable counts and complete word lists and (2) to produce a regression equation which could be used to determine the reading difficulty of the programmed materials.

The sample of materials consisted of three units and the pupil sample consisted of 559 elementary school children in grades four, five, and six.

An error count was used as the measure of the criterion variable and the variable was validated by correlating the counts on this variable with five measures of achievement and an IQ score. The correlation coefficients were all negative (as predicted) and significant at the 0.01 level.

The independent variables were determined by an analysis of the literature and the programmed science materials. Counts on these variables were obtained by using a computer. Seventeen independent variables were chosen to be used in the initial regression equation to predict the dependent variable. A regression routine for the computer was used to obtain the equations from the sampling procedures and deletions were made to determine which independent variables contributed significantly to the final equation. Null hypotheses were stated and checked for each of the seventeen independent variables, and for the significance of difference between the multiple R's using different sampling procedures: (1) random sampling; (2) sequential sampling; (3) cluster-sequential sampling; and (4) total sampling, on the established mathematical regression models.

Only three of these variables were found to contribute significantly to the regression equation for predicting the criterion variable: (1) the average number of mathematical and scientific words per sample; (2) the per cent of frames that were response frames per sample, and (3) the average number of review frames per sample, respectively. The final regression equation had a multiple R. of 0.502.

9:00 to 9:55 A.M. SESSION 1.6 DIVISION C

CLASSROOM INTERACTION I

CHAIRMAN:
Harold Berlak, Washington University

PAPERS:

The Development and Evaluation of Techniques to Assess Teacher Directiveness

BRUCE W. TUCKMAN, Rutgers—The State University

The purpose of this investigation was to develop and validate a measure of *teacher directiveness* for use with students.

An extensive literature search was undertaken to identify scales that have already been built to measure the teaching style which we have labelled "directiveness." The results of this search encouraged our belief that students were reliable judges of teacher behavior, but led us to the decision of developing our own scale since those in existence failed to separate directiveness and negative affect (i.e., warmth-coldness).

Starting out with a detailed operational definition of directive teaching, three scales were constructed: (a) Student Perception of Teacher Style (SPOTS), (b) Observer Rating Scale (ORS), (c) Teacher Style Checklist (also an observer measure but differing in format from (b).

Twenty-two teachers in two vocational high schools were observed by two observers who each completed both observer measures for each teacher. All of the students in a single class of each teacher completed the SPOTS and each teacher completed a personality measure, developed by Tuckman, called the Interpersonal Topical Inventory (ITI), which is a measure of personality structure ranging from concrete-dependent to abstract-independent.

The judgments of the two observers on each of the rating scales had a corrected reliability of about .80 (the observers only observed each teacher once); the ratings were subsequently pooled to form a consensus judgment, on each scale. The agreement between the two observer measures was high ($r = .78$). Furthermore, a strong relationship was obtained between the Checklist score and the ITI; abstractness score ($r = .59$) indicating that the more abstract a teacher was, the more *non-directive* he was. This finding is consistent with others that have been recently reported in the literature.

The relationship between the student measure and observer measures were not as high as expected. The more accurate of the two observations on the Observer Rating Scale correlated .53 with the mean judgments on the SPOTS. Moreover, of the 22 teachers observed, the ORS and SPOTS would lead to the same classification, i.e., directive or non-directive, on 15 of them.

Extensive internal analyses were undertaken on the student measure. The SPOTS was found to be reliable, i.e., the correlation between the tenth most accurate student and the total score across all students was .69 indicating high inter-student agreement. Item-whole correlations were significant for all but eight of the 32 SPOTS items. Finally, a factor analysis indicated a strong 14 item factor which appeared to be "pure" directiveness with other smaller groupings having less meaning.

◆ ◆ ◆

The Effect of Teacher Use of Student Ideas on Student Verbal Initiation

EDMUND T. EMMER, The University of Texas

To determine whether increased teacher use of student ideas would produce increased student verbal initiation, an experiment was conducted in sixteen second grade classes. Flanders' Interaction Analysis categories were used to code teacher and student talk. Teacher statements which clarify, expand upon, accept or make use of student ideas are coded as *teacher use of student ideas;* pupil talk which is voluntary, going beyond the teacher's requirements, is coded as *student initiation.*

There were three stages in the experiment:

I. Baseline observations of classroom interaction were made for five half days, under normal conditions.

II. Observations were made during two 20-minute periods in which the teachers attempted to elicit student initiation in any way they felt was appropriate.

III. Observations were made during two more 20-minute periods in which the teachers attempted to elicit student initiation. One group of teachers was asked to attempt specifically to use student ideas more often; the other group was not given this instruction.

Topic and the sequence of presentation of the topics were randomized during II and III.

The results indicate that teachers who increased their use of student ideas during II and III, as compared to the Baseline observations in I, elicited a significant increase in student initiation. Also, teachers who increased their use of student ideas in III, compared to II (this comparison being controlled for topic and sequence of presentation), elicited a significant increase in student initiation. However, teachers who decreased or remained stable in their use of student ideas did not elicit an increase in student initiation. Changes in student initiation were unrelated to other types of teacher talk (e.g., praise, asking questions).

It was concluded that increased teacher use of student ideas produced increased student initiation.

◆ ◆ ◆

Relationships Between Teacher Behavior and Pupil Creativity in the Elementary School

WILFORD A. WEBER, Syracuse University

This study, based on the view that the fostering of the creative potential of pupils is a valued educational goal, investigated the relationships between specified patterns of teacher behavior and pupil creativity. A conceptual model—suggested by the theoretical speculation of Carl Rogers and the research of Ned Flanders—led to the testing of hypotheses relating the teacher behaviors which 180 elementary school pupils experienced during a four year period (identified as either indirect or direct through use of the Flanders System of Interaction Analysis) and the creativity of these pupils (as measured by the Torrance Tests of Creative Thinking).

Analyses of variance indicated that pupils who had experienced indirect teaching during their first three years of school achieved significantly higher scores on each of the three measures of verbal creativity—fluency, flexibility, and originality—than did pupils who had experienced direct teaching during their first three years of school ($p < .01$). In the main, however, the fourth year experience had little, if any, influence on the verbal creativity scores. Analysis of the data from the four measures of figural creativity—fluency, flexibility, originality, and elaboration—yielded less clear-cut findings. Unlike the results involving the measures of verbal creativity, differences due to the type of teaching pupils experienced during their first three years of school were minimal (the highest F ratio being only 1.14). Findings with regard to the fourth year experience were inconsistent.

On the basis of these results, it was concluded that certain aspects of the creative potential of pupils—particularly verbal potentialities—are fostered more under the influence of indirect patterns of teaching behavior than they are under the influence of direct patterns of teaching behaviors. This study, therefore, lends additional support to the mounting evidence which has indicated the positive effect of indirect teacher behavior.

◆ ◆ ◆

Optimum Teacher-Pupil Interaction for Pupil Growth

ROBERT S. SOAR, University of Florida

Previous work by the investigator relating observed classroom behavior to pupil growth produced results in conflict with those of other studies.

Specifically, indirect teacher behavior (as measured by the Flanders' System of Interaction Analysis) failed to relate to pupil subject-matter growth, although it did relate to growth in creativity. Additional data indicated that the teachers studied were highly indirect on the average. This suggested non-linear relations at the extreme for subject-matter growth, and an optimal level of teacher indirectness beyond which greater indirectness produced less pupil growth. The findings for creativity suggested that a higher degree of indirectness, perhaps higher than the values observed, was optimal for it.

The proposed paper would present a reanalysis of the data to test these hypotheses, and to extend them to a measure of emotional climate.

Hypotheses:

1. The relations between the dimensions of teacher control and emotional climate on the one hand, and pupil growth on the other, are, at the extremes, non-linear. Consequently, an optimum level of the behavior exists which differs from the extreme.

2. The level of classroom behavior which is optimal will shift systematically with the abstractness or divergence of the measure of growth.

Procedure: Residual true gain means for vocabulary, reading, and creativity were calculated for 55 elementary classrooms. Classroom behavior was measured by Flanders' Interaction Analysis, and an OScAR-type schedule, and reduced to factor scores. The relation of a factor reflecting teacher control, and another reflecting emotional climate with each of the sets of true gain means, was examined by polynomial regression. The differences in the curves produced conform generally to expectation, by inspection; shifts of the optimum levels will be tested by the McNemar test of significant change.

9:00 to 9:55 A.M. SESSION 1.7 DIVISION D

ATTITUDES TOWARD EDUCATION

CHAIRMAN:

Winton H. Manning, College Entrance Examination Board

PAPERS:

Dimensionality of Attitudes Toward Services of State Departments of Education

W. L. BASHAW, JAMES B. KENNY, WILLIAM LANDRUM, and R. ROBERT RENTZ University of Georgia; and FOSTER WATKINS, Auburn University

The structure of attitudes toward services rendered by state departments was studied in six southeastern states with a 70 item scale. The instrument was developed from a pool of 180 items. The development procedures will be described fully.

It was administered to randomly chosen samples of superintendents (N = 671), central office personnel (N = 404), principals (N = 627), teachers (N = 3,634), and other local personnel (N = 373).

Data from each group were factored and studied for dimensionality. Factors were compared across groups and identified. Dimensionality was studied by graphing root size and root differences (Cattell's scree criterion). The graphs are presented and discussed.

Factor similarity indexes were used to match factors. Corresponding factors had similarity indexes in the range of .83 to .98. Two factors were common to all data and a third factor was common to three sets of data.

The third factor dealt with university-state department relationships and was not found in teacher or other local data. Evidently knowledge of university-state department relations was insufficient at the local school level to allow systematic responses to this subset of items.

The two factors common to all groups consisted, respectively, of positively and negatively worded items despite the fact that the instrument was developed in a way intended to minimize response set factors. A partial explanation is that the items were all unidirectional, not bipolar. The two factors correlated from .0 to —.7, depending on the group. The two factors correspond, to a degree, to Kerlinger's Conservative and Progressive factors. Kerlinger's factors also were characterized by a split between negatively worded items and other items.

The results are compared to those of Kerlinger, Hines and other local unpublished data. Our results are equivocal with regard to accepting Kerlinger's two-orthogonal-factor second order solution or Hines' one-factor solution.

Measurement of Attitudes in Three Groups of Elementary School Pupils Using a Semantic Differential Scale

CLAUDIA H. HAMPTON and ROBERT W. PATTON, Jr.* Los Angeles City Schools

It has been suggested by the literature that culturally disadvantaged pupils have negative attitudes toward education which are reflected in a lack of academic achievement and inability to take maximum advantage of the school experience. The implication is that sound educational planning should give special consideration to those factors deemed to have direct influence on attitude formation.

If it can be determined that there are differences in attitude which correlate with differences in achievement, it might be possible to identify those factors which tend to favor positive attitudinal development. The purposes of this study were to investigate the attitudes of selected sixth grade pupils presumed to be educationally disadvantaged toward a number of concepts related to school learning and to determine whether there were attitudinal differences associated with differences in school achievement. A semantic differential instrument was developed based on the evaluative factor only and consisting of eight concepts arbitrarily selected for their relevancy to the school situation. A five point rating scale was used in accordance with Osgood's suggestion that this adaptation be made for use with children. This instrument was administered to three sixth grade classes in an inner city school. Teachers assessed achievement levels by assigning a single letter grade to each individual pupil in their class; cumulative records were examined for previous grades and on the basis of these sources of information the pupils were divided into three groups—high, average, and low achievers. Data from a random sample of seventeen tests from each group were analyzed. Mean ratings were calculated for the three groups and an analysis of variance was performed to determine the significance of differences between factor scores. No significant differences in any of the factor scores were found. The pupils studied expressed generally favorable attitudes toward school and school-related concepts which, if they represented true feelings, is in contrast to some statements found in the literature.

◆ ◆ ◆

Measuring Attitude Toward Educational Change and Innovation

GEORGE SCHLESSER, ROBERT CROWLEY, DAVID JENKINS, AND PETER FAY, Colgate University

This research project attempts to design and develop an instrument to measure attitude toward introduction of educational innovations as part of the Regional Colgate-School Research and Development Study sponsored by the Charles F. Kettering Foundation. The instrument is one criterion measure in an action phase to introduce effectively new curricula and technologies. The method used is basically the Thurstone equal appearing interval type using the Likert method of summated ratings at one stage. Action is to include effective introduction of innovations in 26 schools.

Eighty trial items were written representing varying degrees of favorable and unfavorable attitudes toward innovation. These were submitted to 102 volunteers to rate on a scale from 1 to 9 as to the degree of favorableness each item seemed to represent.

Three preliminary forms of 22 items each were prepared consisting of scaled items chosen for lack of ambiguity, Q less than 1.00. Three shorter forms of ten items each were then constructed by selecting items which had higher discriminating power.

Q values ranged from .28 to 1.63. All items with Q value over .99 were rejected as being subject to variable interpretation. Inter-form reliability for the three 22 item editions in homogeneous age groups averaged .76 (.83 for two best forms). Interform reliability for the three 10 item editions was .59 (.79 for best two forms). These reliabilities are sufficient for group comparisons.

Two preliminary studies have been made using short form with 47 chief administrators and with 26 secondary school teachers. These studies show that the scale has promise with respect to construct validity. Younger administrators tend to indicate a more favorable attitude toward introducing innovations than middle age (.10 level). Also, more administrators measured as having favorable attitudes toward innovations "very definitely wished to participate" in the action research phase of the project (.05 level). Comparable results were found for the teacher group as shown by Chi Square's.

◆ ◆ ◆

The Construction and Validation of a Scale to Measure Children's School Morale

LAWRENCE S. WRIGHTSMAN, RONALD NELSON, and MARIA TARANTO, George Peabody College for Teachers

The purpose was to conceptualize salient aspects of children's school morale and to construct a multi-dimensional scale adequate to measure it. After a review of the literature, school morale was conceptualized as having seven dimensions: morale about school plant, administration-regulations, teacher-student relationships, other students, community support and parental involvement, instruction and instructional materials, and general school feeling. An 84-item Likert scale was constructed (12 items for each above subscale) and item-analyzed.

The scale was standardized on 350 5th graders, 350 6th graders, 500 7th graders, 500 8th graders and 500 9th graders in three school systems in Tennessee. Sex differences, school differences, and grade-level differences emerged on most subscales and on the total score.

Girls had significantly better morale at all grade levels than did boys. Morale increases from the fourth to fifth grades; it drops from seventh to eighth to ninth. Scale scores had only low to moderate relationships with achievement and aptitude measures.

The scale was validated by comparing the students' scores with teacher nominations of students with good or poor morale. Differences were generally significant. The scale appears to be a useful device to evaluate the effects of E.S.E.A. Title III programs.

9:00 to 9:55 A.M. **SESSION 1.8** **DIVISION C**

VERBAL BEHAVIOR I

CHAIRMAN:

Joan L. Prentice, Indiana University

PAPERS:

The Effects of Syntactical Violation, Mode of Presentation and Practice on the Free Recall of Verbal Materials

LAWRENCE W. LEZOTTE and JOE L. BYERS, Michigan State University

Previous research has shown rather unequivocally that syntactical manipulation of sentences effects the recall of those sentences. The objective of this experiment was to determine whether the amount of free recall differs when sentences of varying degrees of syntactical violation were presented to fifth grade boys and girls by an auditory, visual, or auditory and visual mode of presentation.

To examine this problem each of fifty-four boys and girls were presented stimulus materials by one of the three modes of presentation. Three levels of syntactical violation were employed:

1) four six-word sentences presented in a correct syntactical order;

2) anagrams made up of four six-word sentences each of which had been randomly scrambled to violate the correct syntactical order; and

3) random strings of words, also in six-word units, which could not be transformed into either sentences or phrases.

To insure stimulus uniformity the seventy-two words were equated for frequency of usage. Initially, twelve syntactically correct sentences were generated and randomly assigned to a stimulus condition.

The three types of stimulus materials were presented to each subject six times (trials). After each trial the subjects were asked to recall and record as many of the stimulus words as possible. The number of correct responses to each type of stimulus material was analyzed by a multiple factorial analysis of variance procedure with repeated measures.

Girls were found to have recalled significantly more words than boys over the entire task, and across all list types and trials. The subjects from the elementary school known to have significantly lower measured educational achievement were also found to have recalled significantly fewer words on each trial and across all list types. Mode of presentation had no significant effect on the recall of the verbal materials.

Syntactical violation was found significantly to effect recall; however, words presented in a semi-grammatical form were recalled less well than words totally lacking in grammaticalness. This finding is significant in that it contradicts the earlier literature which reported findings using college-age subjects. This apparent contradiction will be discussed in terms of the problem solving nature of the experimental task.

◆ ◆ ◆

Verbal Organization and the Facilitation of Serial Learning

JOEL R. LEVIN and WILLIAM D. ROHWER, Jr., University of California, Berkeley

The guiding hypothesis for the present study is that verbal organization is relevant to serial learning but only if the type of organization is consistent with what is ordinarily learned when a serial list is acquired.

Previous research has shown that no facilitation occurs in serial learning when subjects are required to incorporate pairs of successive items into a sentence. However, the kind of verbal organization used, that is, successive discrete sentences, is not consistent with what is presumably acquired in serial learning. A different type of verbal organization, specifically, one that confers on all items in the serial list membership in a single unit, would be expected to produce facilitation not accountable in terms of enhanced response learning. The present experiment was designed to test this prediction.

96 fourth- and fifth-grade children participated in a serial learning task. Subjects were required to learn a 14-item list of common English nouns in either the traditional manner, or with the aid of individual phrases, or in the context of a single sentence.

The dependent variable was the number of nouns correctly recalled over four anticipation trials. It was found that subjects who were given the single unifying structure, the sentence, performed significantly better than did subjects who learned under non-unified conditions. Subjects who were given the nouns in the context of individual phrases performed no better than those who learned the list in the traditional manner, thus corroborating the data of the earlier experiment cited.

The results of the present study support the initiating hypothesis, namely that verbal organization is relevant to the integration of a sequentially ordered set of responses. A response integration interpretation of serial learning is discussed, and problems worthy of further investigation are presented.

◆ ◆ ◆

The Relationships among Reading Ability, Grade Level, and Syntactical Mediation in Paired-Associate Learning

CARIN S. COOPER, University of Wisconsin (Wayne Otto)

The purpose of the present study was to examine the relationships among syntactical mediation, reading ability, and paired-associate (PA) learning.

It was hypothesized that good and poor readers would differ in the number of trials needed to learn a PA list, and that mediation instructions would facilitate learning. 64 Ss were used in a 2 x 2 x 2 x 2 factorial design with two reading levels (good and poor), two grade levels (second and fourth), two types of instructions (mediation and non-mediation), and two sexes. *Ss* were chosen on the basis of consistently high or low reading achievement scores and average intelligence test scores. All Ss learned an eight-pair list of pictures of common objects to a criterion of one errorless trial or 20 trials, and then they were asked how they had been trying to remember.

The results showed that the type of instructions received most affected the speed of learning, with the mediation instructions resulting in fewer trials. The older Ss also learned in fewer trials. A significant interaction between reading ability, grade, and instructions was discussed in terms of IQ and spontaneous mediation. The fourth grade good readers had higher IQs and mediated spontaneously more than the fourth grade poor readers.

◆ ◆ ◆

Verbal Facilitation of Paired-Associate Learning in Children and Adults

NANCY SUZUKI and WILLIAM D. ROHWER, Jr., University of California, Berkeley

Children learn noun pairs faster when the pairs are presented in noun-verb-noun sequences than when presented in noun-conjunction-noun sequences (Type I strings). The problem is to determine whether the superiority of the verb strings would still be observed even if both kinds of strings are lodged in complete sentences, of the forms, noun-verb-noun-conjunction-pronoun and noun-conjunction-noun-verb-pronoun, respectively (Type II strings). Auxiliary question is whether the position of the verb and conjunction connectives within the Type II sentences would affect performance. The purpose of the present experiment was, first, to test the hypothesis that performance in children would vary as a function of connective form class regardless of string type or sentence position, and secondly, to determine whether or not the results would differ depending on the age of the *Ss*.

The design was a 2 x 2 x 2 factorial in which the principal variables were Age, Connective Form Class and String Type, with Trials and Sentence Position as repeated measures factors. 40 fifth-grade children and 40 college students were randomly assigned to the four experimental conditions. The learning materials were 24 pairs of nouns, and all *Ss* were given 2 complete study-test trials.

Analysis of the data revealed a difference in performance as a function of Connective Form Class regardless of String Type or Sentence Position in the fifth-grade sample. A difference in performance was associated with String Type in the adult sample.

9:00 to 9:55 A.M. SESSION 1.10 DIVISION A

PLANNING AND DECISION-MAKING IN EDUCATIONAL ORGANIZATION

CHAIRMAN:

John Sokol, Region II, United States Office of Education

PAPERS:

Portland's Primary PERT Project

JAMES N. HOLMES, Supervisor of Testing and Research for the Portland Public Schools

PERT is an acronym for Program Evaluation and Review Technique. It is a concept which emphasizes the importance of the reciprocal relationship between information and adaptation. The use of this concept by the Navy's Special Projects Office in the development of the Polaris missile resulted in a spectacularly successful program which enabled the first Polaris missile to be ready nearly three years ahead of schedule. The key to the Navy's successful program was its development of a system to gather and process information which improved the decision-making process of the program.

The *purpose* of the Portland Primary PERT Project is to increase the effectiveness of compensatory education in Portland's underachieving schools. The *procedures* include the development of a system for the collection and processing of information which will enable school personnel to adapt their methods to be more appropriate for the needs of pupils. A series of seven locally developed reading diagnostic exercises will be administered each month to approximately 600 primary level children in the experimental group and 200 children in the central group. The data will be processed by a computer and will be reviewed by teachers and administrators to determine what actions could be taken to improve the project. Each month the records will be graphed to show present achievement, past progress, and future needs.

The *results* and *conclusions* will be available by February but the emphasis of the paper to be presented would be upon conceptual and procedural considerations. Many if not all urban school systems are becoming increasingly concerned with ways to improve organizational communications to resolve problems and attain objectives. The basic concept underlying PERT, along with certain aspects of information theory and computer technology, may provide some new insights for educational leaders who are looking for ways to improve the effectiveness of their educational enterprises.

The Use of Ordering Principles in Responding to Decision Situations

GERALD R. SMITH and JOSEPH FERREIRA, Syracuse University

This study examined the relationships represented by the following questions:

1. Do decision makers tend to respond to situations they perceive as more important before they respond to those they perceive as less important (or vice versa)?

2. Do decision makers tend to respond to situations they perceive as more complex before they respond to those they perceive as less complex (or vice versa)?

3. Are age, length of administrative experience, type of training, and other such variables associated with differences in order of response, order of importance, and estimates of complexity?

The rationale for this study is that decision makers have a decision tolerance level which limits the number of decision situations with which they are willing to be confronted at any one time. While the decision tolerance may vary greatly from one individual to the next, each has an upper limit of endurance. When this limit is approached, the individual will put into operation a number of techniques to reduce the number of decision situations facing him. The number and variety of such techniques are undoubtedly large. This study assumes that dealing with the simpler and less important situations are two of the most frequent ways of reducing the number of decision situations.

The twelve items of the Midville High School In-basket were administered to a group of 125 subjects with varying personal, educational, and experiential backgrounds. The subjects were requested (1) to record the order in which they would respond to the in-basket items in an actual school situation, (2) to estimate the time required for them to take both initial and terminal action on each item (a measure of complexity), and (3) to rank the in-basket items in order of their importance. Importance was to be judged on the basis of the seriousness of the likely consequences for the individuals involved or the organization. Correlations were computed between the order of response and the time estimates and the order of response and order of importance. The chi-square test for independent samples was used to determine whether statistically significant differences existed in the responses of groups differing in age, length of administrative experience, type of training, and other such variables.

On the basis of an incomplete analysis of the data, the following results were obtained. There is a high positive relationship ($r = .6188$) between the order of response and the importance attached to the situation. Thus, subjects tended to respond first to those situations they perceived as most important. This finding is contrary to what was expected in view of the rationale suggested earlier. Two explanations appear possible. First, the subjects may have given what appeared to be the most professionally acceptable answer: that is, they may have replied as they thought they were supposed to. Second, the concept of importance may be confounded with the concept of urgency where the need for immediate response would be more expected.

There is a moderate positive correlation ($r = .3340$) between order of response and the brevity of initial time estimates. In other words, subjects do tend to respond first to the simpler situations and put off responding to the more complex ones. This finding is in keeping with the basic rationale of the study. Finally, none of the characteristics of decision makers was associated with differences in order of response, order of importance, or estimates of complexity.

◆ ◆ ◆

Game Theory Analysis of Simulated Collective Negotiations

GEORGE H. ROSS, College of Education, University of Iowa

At a symposium on *Feedback in Simulation Techniques* at the annual meeting of AERA in 1967, Horvat stated: "There is some doubt that games which involve simulation of real-world situations can ever be subjected to the rigors of game theory analysis." This paper is an investigation into the use of negotiated games, a branch of game theory, to analyze a simulation of collective negotiations based on the Negotiations Game developed by the University Council for Educational Administration.

School board members and teacher organization representatives from Iowa's fifty largest school systems were invited to participate in this simulation. Prior to the actual negotiations session each team of board negotiators and each team of teacher negotiators assigned utility values to each of the issues. Based on these values an optimum solution or agreement was arrived at through game theory for each group of negotiators. These optimum solutions were compared with final agreements that the groups arrived at by playing the Negotiations Game.

The paper will report the use of analysis and the results of its application to these negotiations sessions.

◆ ◆ ◆

Position Analysis for Administrative Personnel in Public School Systems

LEO GERALD MAHONEY, University of Houston

The problem was to *adapt* techniques of position analysis for use with administrative personnel in public school systems. This study has been primarily concerned with the identification of the functions and responsibilities of school administrative personnel, and with the development of an adequate and appropriate conceptual framework for analysis of their positions.

Materials and techniques of personnel management in business and industry, government, and education have been drawn upon: 1) to identify the components and the analysis procedures of organization planning, staffing, and wise compensation structure; and 2) to apply and adapt procedures of position description, position analysis, and position evaluation for use with school administrative personnel. Carefully selected and adapted samples and illustrations have been provided, culminating in a model plan for position evaluation employing a combination of the factor comparison and point systems of position description and analysis.

The field of personnel management is highly specialized. Detailed plans and operational details were found to be classified information; therefore, the working materials contained in this study constitute a unique resource for use in positional analysis with school administrative personnel.

1. Plans and procedures of positional analysis should be adapted, rather than adopted.

2. Defining and grouping administrative activities and establishing authority relationships among them is necessary to accomplish specified organizational and operational results.

3. Position description questionnaires provide the practical starting place for analysis of professional positions.

4. Performance appraisals should be carried on in terms of planned, workable, and rational procedures.

5. Wise compensation structures can be best determined through evaluation and appraisal within the framework of an explicit, systematic plan.

6. The model evaluation plan presented in the study is applicable to most school systems, when appropriately adapted to local situations.

7. Systematic evaluation of school administrative personnel requires the services of a specialist-consultant.

8. Competent school administrators can be found and/or developed, but first they have to be identified by a performance appraisal, then trained for and assigned to positions that are clearly described and for which they are qualified.

The recommendations offered were:

School administrators should re-examine organization structures and personnel practices with the aim of introducing personnel procedures adaptable to, and commensurate with the needs and demands of, their executive and managerial functions and responsibilities.

Training stations or centers, on a local or regional basis, should be established for the purpose of providing school administrative personnel supervised access to practical applications of position analysis in their school systems.

Universities responsible for preparation of school administrative personnel should provide substantial and operational training in position analysis concepts and techniques.

Centers for research and experimentation in this field should be established, preferably on the basis of cooperative arrangements between qualified universities and interested public school systems.

Specialists in school personnel management and analysis should be used by school systems as consultants with their administrative personnel, at least until the administrative personnel have reached the point of managing the processes involved with competent safety and confidence.

9:00 to 9:55 A.M. **SESSION 1.11** **DIVISION A**

POLITICAL ASPECTS OF EDUCATIONAL POLICY DETERMINATION

CHAIRMAN:

Michael L. Thomas, University of Texas

PAPERS:

Legislative Subalterns and Policy Process Control

HAROLD I. GOODWIN, West Virginia University

The policy roles of certain categories of legislative internals remain but dimly understood. This conceptual analysis of an interview study of a broad spectrum of legislative policy participants in the California State Legislature explores the bases of control over the policy process exercised by one class of internals, legislative subalterns.

Two major control bases emerged: the formal basis which structures subaltern participation, and the operational basis upon which turns the working control over policy formation. The two dimensions interact complexly.

The acquisition of political power by legislators, based upon their formal organizational roles and upon acquired peer status, results in a power hierarchy among legislators. This differential orders, in part, the relationships of subalterns, who accord positions among themselves accordingly. Secondly, legislators make differing authority allocations to subalterns for involvement in and control over education legislation formulation. These specifications of latitudes or constraints condition the interaction patterns among subalterns vis-a-vis education legislation and over access of externals to the policy process.

More importantly, subaltern control hinges, first, on their independently acquired peer status grounded in a demonstrated combination of policy content expertise and political maneuver. Second, control is reflected in real or potential indebtedness among subalterns, relevant legislators, and between external interest figures and a subaltern. Third, control is reflected in the subaltern's capacity to counter another's goal aspirations while functioning within his legislator's performance expectations.

The interaction of formal and operational controls establishes a subaltern power network structuring the subaltern control exercised over policy processes. Disregarding subalterns handicaps subsequent opportunities to infuse values into education legislation. While educators are generally unaware of the subaltern power network and its policy process implications, persons understanding and privy to the power network exercise a disproportional influence over the policy process involving education legislation.

Issues and Influentials: The Decisional Process in School District Reorganization

GLENN E. WILEY, Illinois State University

The citizens in nine rural school districts, when faced with the prospect of rapid industrialization, ended a twenty-year debate, and almost overnight reorganized their districts to form the Delta School District. The purpose of this study was to (1) identify the most influential individuals in the Delta School District, (2) determine the action taken by each in the decisional process for school district reorganization, and (3) analyze these actions to determine at what stage in the decisional process each influential participated, and the type of influence he exerted.

The procedure used in this study was an adaptation of the "reputational approach" as introduced by Floyd Hunter, supplemented by an analysis of the Delta School District reorganization process. Data were collected from public records, newspaper reports, and personal interviews. The major findings of the study reveal that in the Delta School District:

1) school leaders and association heads are equally capable of identifying influential individuals;

2) school leaders are able to identify those individuals who are most influential in bringing about school district reorganization;

3) there is a monolithic power structure;

4) the fact that five men were perceived to be influential in all areas of community decision-making suggests that a "pyramid of power" existed;

5) retaining control over financial resources was the prime consideration which brought the initiation of school district reorganization;

6) *public leaders* were the most influential individuals in initiating school district reorganization;

7) public officials are most often the *public leaders;*

8) there were no influential individuals who at the time of reorganization operated "behind the scenes" to control the decision.

It was determined that the use of the paradigm of decision-making described in this study was a suitable method for analyzing the decisional process involved in the formation of the Delta School District.

◆ ◆ ◆

Effects of Legislation and State Board Rulings upon Teacher Supply

DUKE B. HUBBARD, University of California at Claremont (Conrad Briner)

It was the purpose of this study to investigate the effects of legislation, since 1961, and State Board of Education rulings upon the supply of elementary school teachers within the State of California. A survey was undertaken involving 1136 elementary and unified school districts and 46 accredited elementary teacher training institutions. The survey sought to ascertain whether school districts had encountered increased difficulties in the recruitment of new teachers, whether colleges and universities had witnessed an increasing decline in the enrollment of elementary education candidates since 1961, and whether or not difficulties arose which could be attributed directly to recent legislation. Non-parametric statistical tests were utilized to analyze the variables involved. Seventy-one per cent of the responding school districts reported increased difficulties in recruitment and 67% of the colleges and universities reported an increasing decline in elementary education enrollment since 1961. School districts stated that of primary importance in their increased difficulties in recruiting new teachers was the decrease in the number of elementary teachers produced by nearby colleges and universities. Local factors such as lack of competition of salary schedule, lack of cultural benefits, and geographical isolation were not thought to be prime causes of district shortages. The two main reasons indicated by colleges and universities for the decline in elementary education enrollment, (1) academic major requirement causing many prospective elementary school teachers to choose secondary education and (2) five years of preparation instead of four, were both consequences of the Fisher Bill.

The results are discussed in terms of the issues involved.

◆ ◆ ◆

The Differential Impact of 1963 and 1965 Indiana State Tuition Support Programs on Financing Indiana School Corporations

DON CALEB PATTON, Purdue University

The major purpose of this study was to determine whether larger and/or wealthier Indiana school corporations were receiving preferential treatment in financial aids received from the State under the 1965 legislative amendments to the State Recognized Program for Instructional Salaries. A subsidiary concern was to analyze interaction effects of two 1965 changes to the 1963 formula for allocating funds to public schools, as these changes tended to relate to factors identified with a general systems model for financing education.

The population was defined as Indiana public school corporations for a period of time that the 1965 plan for allocating funds to public schools, or a similar plan, is applicable. The sample consisted of 417 Indiana school corporations eligible to receive state aid during the 1965-66 school year. Data for the study were obtained from the Indiana State Department of Public Instruction.

Size of school corporation, wealth of school corporation, and allocation formula treatments served as independent variables. A 4 x 4 x 4 factorial analysis of variance design with repeated measures on the third variable (formula treatments) was employed. Size of school corporation was measured by number of pupils in average daily attendance for the 1965-66 school year, who resided in and attended schools in the school corporation. Wealth was measured by the adjusted assessed valuation per resident pupil in average daily attendance for the 1965-66 school year. Four allocation formula treatments were applied to each of the 16 size and wealth categories. Those treatments were:

1) The 1963 Indiana Minimum Foundation Program for Instructional Salaries.

2) The 1963 Program with financial allocation increased from $175 to $185 per pupil.

3) The 1963 Program with an increased incentive aid based on teacher training and experience.

4) The 1965 Recognized Program for Instructional Salaries which was the 1963 formula with both legislated formula changes.

An unweighted means procedure with harmonic n was used to adjust for unequal cell frequencies.

Main effects, interaction effects, and simple main effects where appropriate, were tested by F tests with alpha $= .05$. Where simple main effects were significant, the Neuman-Keuls procedure was used to test for differences between means with $Q_r = .05$.

The major findings permitted the following conclusions to be drawn:

1) Larger and wealthier Indiana public school corporations were not receiving a disproportionate share of state funds allocated to public school corporations for instructional salaries.

2) The two changes in the 1963 Indiana formula for allocation of state aid to public school corporations did not interact to produce significant differential impact on school corporations.

3) Wealth of school corporation was inversely related to amount of per pupil state aid received for instructional salaries in the four corporation wealth categories of this investigation.

4) Variations indicative of inequities in allocating funds were noted for size categories of school corporations.

5) Inequities found in the 1963 formula for allocating funds to school corporations for instructional salaries were not significantly changed by the 1965 amendments to the formula.

10:00 to 11:55 A.M. SESSION 2.2 DIVISION D

PREDICTING ACADEMIC SUCESS

CHAIRMAN:

Warren W. Willingham, College Entrance Examination Board

PAPERS:

Predicting Computer Programming Performance

WILLIAM A. MEHRENS, JOHN F. VINSONHALER, ROGER BAUER*, Michigan State University

One of the fastest growing and most important of the new technical profession is computer programming. Over 50,000 programmers are needed immediately. An additional 100,000 will be needed by 1970. Therefore, it is crucial that we be able to identify potentially good computer programmers. Currently, the literature reveals no batteries designed explicitly for this purpose. It was the purpose of this study to perform some preliminary research for the development of such a battery comparing the predictive effectiveness of various existing tests singly and in combinations.

Sixty-eight students in an introductory programming course were tested with the Aptitude Test for Programming Personnel (ATPP) and the Strong Vocational Interest Blank (SVIB).

A multiple repression analysis was computed using scores on these tests together with scores on the College Qualification Test (CQT) and previous grade point average. Variables included in the analysis were the three part scores plus the total score from the ATPP, three part scores plus total score for the CQT, and the Computer Programmer and Academic Achievement Scales from the SVIB.

The results show that all independent variables were positively correlated with course grade. Single correlations ranged from a low of .30 (ATPP, Part I) to a high of .68 (GPA). Surprisingly, the specially developed ATPP was not a better predictor than the CQT, a standard scholastic aptitude test. Indeed, excluding GPA, the best single predictor was the numerical subtest of the CQT.

Omitting GPA as an independent variable, the best combination of two scores proved to be the CQT numerical and Part II (Figure sequence) of the ATPP ($R = .61$). The best three score combination was formed by adding the Computer Programmer Scale of the SVIB ($R = .65$). The best four test combination was obtained by adding CQT: Verbal ($R = .68$). Combinations of more than four scales did not significantly improve the multiple correlation.

Several conclusions may be tentatively drawn with respect to the prediction of programming performance. First, specially developed programming aptitude tests may not necessarily be better predictors than tests of general intelligence. Second, combinations of tests yield better predictions than single tests. Third, the best predictive combinations appear to be subtests with rather diverse item content, as might be expected. Fourth, prediction was significantly improved by the use of an interest scale. In general, the results indicate that further research should be done on the development of a predictive battery for programming performance.

◆ ◆ ◆

Comparative Utility of Educational Background Data on Airmen at Three Ability Levels

CHESTER J. JUDY, Aerospace Medical Division, AFSC

During periods when heavy demands are made on the manpower resources of the nation it is usual to consider, for military service, larger numbers of individuals whose scores on various service tests are relatively low. Thus, during the Korean conflict, nearly 30% of those accepted for military duty were in a Category IV category (i.e., those whose scores on the Armed Forces Qualification Test were in the 10-30 percentile range). In more "normal" times, however, through the operation of a policy in which an attempt is made to minimize training and other costs, relatively few men at lower levels of ability are brought into service, insofar as this must be done. Selection and classification procedures tend to be geared to this policy; relatively little research attention, especially in the Air Force, has been given to unique selection and classification problems having to do with the optimum use of low aptitude personnel.

This research is a part of a larger task in which the purpose is to evaluate a variety of measures which may be appropriate for use in improving the assignment of men at lower levels of ability. Using multiple regression procedures, technical school grades in a first Air Force course as criteria, and data on four groups of airmen divided into high-, middle-, and low-ability subgroups, the predictive utility of preservice educational data, as compared with the best combination of scores from selection tests, was investigated. It was found that the predictive utility of background educational information tends to remain constant across three ability levels whereas there is a marked decline in the usefulness of scores from present tests in going from high- to low-ability subgroups. It was concluded that the preservice educational record represents a promising source of useful information.

◆ ◆ ◆

Ability, Family Socioeconomic Level, and Advanced Education

LYLE F. SCHOENFELDT, American Institutes for Research

The importance of family background as an intervening variable in the relationship between ability and advanced education has long been of interest. However, statistically adjusting for differences in ability among career groups tends to obscure other relationships. The purpose of the present study is to explore the influence of family socioeconomic level on advanced education in a controlled setting: nursing education.

Certification as a registered nurse remains as one of the few educational goals that can be approached by means of two major alternate paths: attendance at (1) a 3 year school of nursing or (2) a college offering a major in nursing. The latter leads to a baccalaureate degree as well as certification, which tends to make it somewhat more prestigious.

The basic sample included girls who participated in the nationwide TALENT 5% probability sample as 12th graders in 1960 and also responded to the 1 year follow-up questionnaire in 1961. From among these 27,000 girls, over 1,000 reported they were studying for an RN degree. The 690 girls who also reported attending a 3 year school of nursing were placed in one group while the 380 attending a 4 year college and majoring in nursing were placed in a second. These two groups were then compared on the maximum performance scores, typical behavior scales, and socioeconomic background items which comprised the Project TALENT two day test battery.

The two groups studying to become registered nurses did not differ significantly on any of the 36 information scales, 22 aptitude and achievement tests, or 27 interest and temperament scales which comprised the Project TALENT test battery. However, there were both significant as well as substantial differences between the two groups on a number of family socioeconomic characteristics. A paradigm is developed illustrating the influence of socioeconomic background on the level of advanced education selected after high school graduation.

◆ ◆ ◆

The Utilization of Factor Scores from Biographical Information to Predict Learning Within a College Course

DAVID D. STARKS, The University of Michigan and JOHN F. FELDHUSEN, Purdue University

The purposes of the present study were to derive a factor structure from biographical information and to assess the predictive value of the factors obtained with respect to two measures of learning.

Data from an 894-response Biographical Inventory (BI) were submitted to the principal axis method of factor analysis and were followed by orthogonal rotations in fourteen subsets of items. Marker variables selected from among the highest loading items on the rotated factors were used in analyses with four, two, and finally one subset of items. The final analysis yielded eleven factors.

The two criteria of learning were a multiple-choice content examination and the Minnesota Teacher Attitude Inventory (MTAI). Multiple regression equations incorporating the eleven BI factor scores, reading ability, and MTAI pretest scores were developed.

Predictions of the two criteria were carried out both within the total sample and within subgroups formed by moderator variables. The moderators were reading ability and attitudes toward teaching at the beginning of the course.

For each of the criteria the BI factors were found to be significant predictors and were also found to produce significant increments in prediction when they were employed in addition to reading test and MTAI pretest scores in the total sample and for the subgroups characterized by high reading or MTAI pretest scores, but not for the subgroups with low reading or MTAI pretest scores.

It was concluded that factor scores derived from biographical information were useful in predicting performance within a course and that the moderator variable aided in prediction by identifying those students for whom the information is more useful.

10:00 to 10:55 A.M. **SESSION 2.5** **DIVISION C**

LEARNING AND INDIVIDUAL DIFFERENCES I

CHAIRMAN:

Richard E. Ripple, Cornell University

PAPERS:

Anxiety and Response Mode to Programmed Instruction

SIGMUND TOBIAS and JOHN H. WILLIAMSON, City College, City University of New York

Prior research has failed to demonstrate any relationship between anxiety and programmed instruction. In these studies anxiety was presumed to be operating on the basis of *S*s scores on anxiety scales. This assumption was subjected to a direct test in the present study by the employment of a no-reinforcement condition which, in the presence of anxiety, should clearly have led to impaired performance of high anxiety subjects.

A total of 104 students, recruited from educational psychology classes, served as *S*s in a 2×3 factorial design. A median split of scores on the MAS defined high and low anxiety groups; and *S*s were then assigned to one of three treatments: two treatments involved constructed response with and without reinforcement, and a third involved a reading, no response, version of the same program. A 90 frame linear program dealing with binary numbers was employed, and achievement was determined on the basis of a post-test. Student familiarity with the subjct matter was controlled by using *S*s' pre-score as a covariate. An attitude questionnaire was also administered.

A 2×3 analysis of covariance of both the post-test scores and the attitudinal data revealed no significant main effects or interactions. The means of contrasting anxiety groups under the no reinforcement condition were also found not to differ. The findings raise serious doubts as to the usefulness of research involving anxiety and programmed instruction, so long as anxiety is defined by a test score. Suggestions for further research involving anxiety induction are made.

◆ ◆ ◆

State Anxiety and Task Difficulty Using CAL Media

HAROLD F. O'NEIL, Jr., CHARLES D. SPIELBERGER, and DUNCAN N. HANSEN, Florida State University

Most studies of the influence of anxiety on learning have used artificial laboratory-learning tasks. In this study, the effect of anxiety on the learning of meaningful mathematical materials by computer-assisted instruction (CAI) was investigated. On the basis of Spielberger's Trait-State Anxiety Theory, it was hypothesized that the performance of high anxious (HA) students would be inferior to that of low anxious (LA) students on a difficult CAI task and that HA and LA students would either not differ or the HA students would be superior on an easy CAI task.

An IBM 1400 series system presented the learning materials, and recorded responses of 29 introductory psychology students. The program contained a difficult section on the field properties of complex numbers and an easy section on compound fractions. The measures of anxiety were scores on the A-state scale of the State-Trait Anxiety Inventory and changes in blood pressure. The A-state scale was given four times—at the beginning of the learning task, after the difficult section, after the easy section, and after a three-minute, post-experimental interview. On each of these occasions, blood pressure was also taken.

Results indicated that both state anxiety scores and blood pressure increased on the difficult CAI material and decreased when the students worked on the easy material. The predicted relation between task difficulty and anxiety level was also significant, i. e., HA students made significantly more errors on the difficult material than LA students. These findings suggest that difficult CAI material arouses anxiety, and that in order to facilitate optimal learning in CAI programs, the students' anxiety level must be taken into account.

◆ ◆ ◆

The Relationship Between Test Anxiety and Children's Need for Memory Support in Problem Solving

JOAN E. SIEBER, Stanford University and LAWRENCE I. KAMEYA, University of Michigan

We describe the rudiments of a research strategy which by-passes some frequently encountered problems of definition and inference in the study of anxiety, then report an experiment which embodies that strategy, and tests the following hypotheses: 1) When information must be remembered before a correct strategy can be formulated, high anxious persons will make more wrong choices than low anxious persons, and will less frequently recognize that a solution is erroneous during VTE. 2) When memory support is provided, the difference between high and low anxious persons in frequency of wrong choices and failure to recognize errors during VTE will be reduced.

40 fifth and sixth graders, matched on test anxiety, defensiveness, sex and I. Q., were divided into two groups, each of which solved Porteus Maze tasks and a marble puzzle, with and without memory support, respectively. Porteus mazes providing memory support were the standard type. Mazes not providing memory support were covered by three overlays, each concealing successively smaller portions, hiding endpoints of several cul-de-sacs, thus forcing *S* to remember details earlier revealed to him. The marble puzzle

(without memory support) consisted on a board with a row of nine holes, containing four black marbles at one end and four white marbles at the other. Solution requires transposing marbles to opposite ends. Only one 24-move sequence results in solution. Discovering the correct solution requires remembering move sequences and their outcomes, except in the memory support version which provides diagrammatic records of *S*s' errors.

Anxiety and Sex and Memory Support ANOVAs were performed on a number of wrong choices and errors recognized during VTE for each task. The obtained significant F ratios for Anxiety, and Anxiety × Memory Support, were interpreted as suggesting that the interference of anxiety with sort-term memory may be partially offset by instruction of anxious children in the use of provided or self-made diagrams or notational systems.

◆ ◆ ◆

Differential Relationships Between Creativity and Intelligence Under Two Conditions of Testing

KENNETH G. O'BRYAN and FREDERIC J. BOERSMA, University of Alberta

In order to investigate further Wallach and Kogan's (1965) suggestion that method of test presentation and frame of reference of the test taker are important variables in the determination of the dimensionality of creativity and intelligence, the effects of two conditions of testing on creativity and its relationship to intelligence were examined. Lorge-Thorndike intelligence tests (verbal and non-verbal) were administered to 46 fourth grade boys under standardized procedures. The Ss were then randomly assigned to two equal groups. Torrance's Unusual Uses (verbal) and Figure Completion (non-verbal) tests of creativity were administered to Group A under conditions similar to those used in intelligence testing. The regular school routine was disturbed as little as possible, and the tests were presented in the context of evaluated, time-controlled tasks. Group B received the same measures outside of the school setting without apparent evaluation or restriction of time. Effort was made to minimize the school atmosphere through the use of uninhibited group play, absence of teacher and authority figures, presence of a highly permissive examiner, and presentation of the tests in an apparently unstructured and spontaneous manner. Group B scored significantly higher ($p < .01$) on both measures of creativity and displayed a substantial reduction in correlation between intelligence and creativity. It was concluded that when emphasis on success orientation and school behaviour patterns are minimized, creativity increases; and that creativity and intelligence possess separate dimensionality. The present study is supportive of Wallach and Kogan's position under a more general definition of creativity using a short term, group-testing procedure.

10:00 to 10:55 A.M. **SESSION 2.6** **DIVISION C**

CONTRIBUTED PAPERS I

CHAIRMAN:

Bruce W. Tuckman, Rutgers, The State University of New Jersey

PAPERS:

Manipulation of Learning Set and Feedback in the Teaching of Drawing

KENNETH R. BEITTEL, The Pennsylvania State University

In two art learning experiments, college undergraduates made sequential drawings over a six and seven week period. They worked in a laboratory studio, singly, one hour weekly, with standardized drawing materials. An electronically operated camera (out of sight) took time-lapse records of each drawing from an overhead front surface mirror. Half the time, there was an assemblage-like, many sided still-life as drawing stimulus; the other half, no stimulus was present.

Both experiments had a balanced, five factor (two levels each) AOV design. Factors A, B, and C were classification (sex, art training, and beginning art strategy). Factors D and E were treatment.

Dependent variables were spontaneity (style), divergency (style), and aesthetic quality, represented as change scores from the benchmark (pretreatment) position. Reliabilities for expert judges averaged .868 over the two experiments.

Effects of *feedback* conditions were slight. In *Experiment I* more *process feedback* aided spontaneous S's and those using mind stimulus, while opposed conditions aided divergent S's and those using the still-life stimulus. In *Experiment II,* direct mediation of *learning feedback* encouraged gains in spontaneity, while indirect mediation increased divergency.

The effects of *learning set* conditions were more pronounced. The still-life increased divergency, the mind condition spontaneity (*Experiment I, implicit learning set*). In *Experiment II,* the *induced learning set* strongly affected style in the direction of stylistic instruction and effects persisted into the post-treatment period.

Apparently strong, *induced learning sets* are necessary to bring about any noticeable change in drawing styles, although subtle effects from *implicit learning sets* can be analyzed. Aesthetic quality changed but little.

◆ ◆ ◆

Phonetic Symbolism in Adult Native Speakers of English

ROBERT TARTE, and LOREN S. BARRITT, The University of Michigan

The purpose was to determine whether adult native speakers of English can agree upon the choice of phonetic labels for visually presented figures.

A forced choice paired-comparison test containing 252 items was made. Each item consisted of 2 figures chosen systematically from a sample of 8 figures. 4 of the figures were circular and 4 triangular. The 4 circular figures varied from a complete circle to an ellipse 1/8 of the height of the original circles with 2 intervening ellipses 1/2 and 1/4 of the height of the original circle. The 4 triangles also varied along the vertical dimension from an equilaterial triangle to one 1/2, one 1/4, and one 1/8 of the height of the original. For all figures the horizontal dimension remained the same. These 8 figures were each paired with every other figure, and presented on slides to subjects. Subjects were asked to decide which of the 2 figures was best labelled with a monosyllabic nonsense sound presented on audio tape. Each of 9 nonsense sounds was individually presented with each pair of figures.

The 9 sounds were derived from a systematic pairing over vowel and consonant dimensions. The 9 sounds presented to S's were:

was, wus, wis; das, dus, dis; kas, kus, kis.

Results show a striking orderliness. There is a consistency within and between S's in their ability to match sounds to figures. A 5 way ANOVA found the interactions between vowel and size, vowel and shape, and consonant and shape all to be significant. ($p > .01$).

The conclusion is that native speakers of English have a sufficiently similar background of experience with language to permit reliably consistent responses on this labelling task. The implications of this finding for future research with samples of "foreign" language speakers are discussed. The use of this technique to identify language universals is suggested.

◆ ◆ ◆

The Effects of Sonic Environment on Individualized Instruction in Learning Difficult and Easy Tasks by High Achievers

EDWARD H. McCARTHY, Bucknell University (Richard C. Anderson, University of Illinois)

When extraneous noises in classrooms become evident to the students engaged in a learning task, the students will either stop or continue with their task. Does their work suffer, and if so how much?

Very little research has been done in education to answer these questions. Therefore, the purpose of this study is to fill the research gap in education by investigating the effects of noise at different decibel levels on difficult and easy learning tasks in a typical classroom setting.

80 sixth grade students having the highest mean scores on the Metropolitan Achievement Test were chosen. They were randomly assigned into decibel groups of easy and difficult learning tasks.

CVC Trigrams were used to eliminate any associations between one item and another. Easy tasks had 70% to 100% meaningfulness. Difficult tasks had 1% to 30% meaningfulness. The CVC Trigrams were presented by paired-associates in a programmed format.

A White Noise at different decibel levels (0, 40, 55, 70, 80) was used to eliminate associations between a noise and a word.

The study was run for three days. The subjects received a schedule and instructions each morning. The noise was playing when the subjects participated. They picked up a learning task, studied it, and then took a test.

The results indicate that performance as a function of the level of activation noise is highest for both difficult and easy tasks at 55 decibels. Performance on both learning tasks increased from 0 to 55 decibels and then decreased from 55 to 85 decibels.

There was a significant difference between 0, 40 and 55 decibel groups and the easy task. There was a significant difference between 55, 70, and 85 decibel groups and the difficult task.

The results indicate that a perfectly quiet room or an excessively noisy room are not the best environmental conditions for students participating in individualized instruction in the sixth grade. Students learn best with a moderate amount of noise.

◆ ◆ ◆

Fatigue and Performance Variability Among Typists

LEONARD J. WEST, City University of New York

To determine (a) fatigue effects of typists at various skill levels of a long, continuous work period at the typewriter and (b) whether consistency of performance increases with skill, 234 typists at skill levels from 5 through 108 words per minute (wpm) typed from ordinary printed prose for 30 continuous minutes in a manner permitting identification of each minute of work and, thereby, scoring for gross wpm and number of errors on a minute-by-minute basis, individually and cumulatively. Variance analyses of speed and error scores among 30 trials (minutes) and six 5-minute blocks of trials used equal-frequency sub-samples of 19 persons from each of nine skill levels (5-14, 15-24, . . . , 85-94, 95-108 wpm), randomly drawn from the original pool of 234 typists. Consistency of performance among nine skill levels was estimated through analysis of variance of V, the Coefficient of Relative Variability, for speed and errors for the 30 trials, using all 234 subjects.

Concerning fatigue effects, highly significant F's were found for 30 and 6 trials for both speed and errors among nine and four skill levels. Significant levels-by-trials interaction was found for speed, but not error, scores. However, absolute speed differences from trial to trial were of no practical consequence (e.g., a .9 wpm decrease between the first and last 5-minute segments). Errors, on the other hand, increased (at a decelerating rate) in successive trials. Concerning consistency of performance, differences in V's among the skill levels were found to be significant for speed and for errors, accompanied by regular increase in consistency with increase in skill for speed, but not for accuracy, of performance.

In agreement with earlier generalizations about comparable perceptual-motor skills, for the typewriting task the consequential decremental effects of a long work period are on accuracy rather than speed of performance. In addition, speed, but not accuracy, shows increasing stability with increase in skill. The former finding, in the light of the small increases in errors from trial to trial, calls into question the conventional instructional practice of restricting work durations to 1 to 3 minutes for the first few months of instruction for reasons of supposed fatigue. Since work durations of any given length have the same effects on errors (but none on speed) at all levels of skill, there is no reason to vary practice durations at different skill levels.

11:00 to 11:55 A.M. **SESSION 3.4** **DIVISION C**

LEARNING FROM PROSE I

CHAIRMAN:

Gerald W. Faust, University of Illinois

PAPERS:

Effective Prose Reading: Shaping and Discriminative Effects of Questions

LAWRENCE T. FRASE, Bell Telephone Laboratories, Murray Hill, New Jersey

This study explored the effect of questions upon the retention of prose material. Ealier research suggests that question pacing, difficulty, and location of relevant material may be important variables in learning from prose.

Subjects were 128 introductory psychology students. A 20 paragraph prose passage from Miller's *Psychology: the Science of Mental Life,* was used. This passage, and retention test questions, had been used in earlier research.

A 2 x 4 x 2 x 2 x 2 factorial design was used. The factors were: 1) question location (before or after paragraphs); 2) question pacing (after every 10, 20, 40, or 50 sentences); 3) content location (question-relevant material located in the first or second part of each 10 sentence paragraph); 4) question mode (multiple-choice or constructed response); and 5) retention items (relevant to questions used in the text or incidental.) The materials were randomly distributed to *Ss* who read through them in a large auditorium and then responded to a 40 item multiple-choice test. Subjects were not allowed to review paragraphs.

Results discussed below were statistically significant at the usual levels.

Findings were consistent with earlier research in that retention was highest with post-questions, and question-relevant content was retained better than incidental. Interactions revealed that retention increased with frequent post-questioning, but decreased with frequent pre-questioning. It was concluded that frequent post-questioning produced effective reading skills while frequent pre-questions interfered with prose structure. Question frequency greater than every 20 sentences produced this decrement. Frequent questioning, either pre- or post-, yielded precise discrimination between relevant and incidental material which took the form of lowered incidental learning without a corresponding increase in relevant learning. This latter finding is consistent with a signal-detection literature in which an information rejection strategy has been hypothesized.

Regardless of the other conditions, retention of incidental material was higher if it followed the question-relevant material in the prose. This finding is also consistent with an information rejection concept of attention.

Question mode (multiple-choice or constructed response had no effect.

◆ ◆ ◆

Retroactive Facilitation in Meaningful Verbal Learning

DAVID P. AUSUBEL, MARY STAGER, A.J.H. GAITE, The Ontario Institute for Studies in Education

The phenomenon of retroactive interference in verbal learning has been clearly demonstrated in many studies which have used nonmeaningful and unconnected materials, chiefly nonsense syllables and paired adjectives. However, there is much doubt as to whether retroactive interference occurs when connected material is meaningfully learned (i.e., when it interacts on a nonarbitrary, substantive basis with established ideas in cognitive structure).

In order to ascertain whether retroactive interference occurs in meaningful verbal learning and retention, the experimental conditions favoring such interference were maximized by using both unfamiliar and conflicting original and interpolated learning materials. The effects of interpolated learning (Buddhism) and of overlearning of the original material (Zen Buddhism) were tested in a 2 x 2 factorial design, using 156 12th grade pupils. Both independent variables significantly facilitated the retention of the original material (overlearning: $p < .01$; interpolation: $p < .05$). The facilitating influence of interpolated learning was attributed to the rehearsal and clarification of the original material which it presumably induced. The absence of a significant interaction term indicated that prior overlearning did not differentially affect the later facilitating effect of interpolation.

These findings, if replicated and given greater generality, would have far-reaching implications for classroom teaching practice. Instead of suggesting (as do the classical retroactive interference findings in the case of rote learning and retention) that teachers scrupulously avoid introducing similar and conflicting material after a typical lesson involving meaningful learning, they imply that such material should be introduced deliberately. This recommendation would be based on the expectation that conflicting interpolated material would encourage the learner to compare related ideas in the original and interpolated sets of material, and thus facilitate retention of the original material through the influence of such intervening variables as rehearsal and clarification.

◆ ◆ ◆

The Effects of Analysis and Evaluation Questions on Various Levels of Achievement

FRANCIS P. HUNKINS, University of Washington

Questions, questioning, inquiry are receiving increasing attention from educational researchers. This particular study had as its objective to discover if a dominant use in social studies text-type materials of analysis and evaluation questions would improve various levels of achievement according to the levels of Bloom's *Taxonomy*. Previous research by this author had determined that such questions did stimulate a significant difference in overall achievement. This research sought to ascertain at what levels this achievement did differ.

Two hundred and sixty sixth-grade pupils in eleven classes were involved in the experiment. Classes were randomly assigned to either experimental Condition A, dominant emphasis (50 per cent) on analysis and evaluation questions, or Condition B, dominant emphasis (90 per cent) on knowledge questions.

During the experimental period of four weeks, specially prepared materials were used by the pupils in each treatment. Condition A materials contained the dominant emphasis of analysis and evaluation questions. Condition B's materials centered primarily on knowledge questions.

At the termination of the experiment, an achievement post test was administered. This test in addition to giving a total score, provided one with six sub-scores, one for each level of Bloom's *Taxonomy:* knowledge, comprehension, application, analysis, synthesis, and evaluation. Analysis of covariance was the principal method utilized in analyzing the subscore data.

Each group of post sub-test scores was subjected to analysis of covariance adjusting for the pre- sub-test scores. Those students experiencing high-level questions, Condition A, achieved significantly better in the areas of application and evaluation than did students experiencing low-level questions, Condition B.

For each of the six sub-score divisions, significant differences were observed between reading levels, with the upper reading levels showing the higher scores.

From the research conducted, the following conclusions are made:

1. The use of high-level questions (analysis and evaluation) produced significantly greater scores in the areas of application and evaluation than did the use of low-level (knowledge) questions.

2. Better readers in both conditions did significantly better in all six areas of the achievement test.

◆ ◆ ◆

Elementary Pupil's Ability to Conceptualize the Main Ideas in Reading

THOMAS C. BARRETT and WAYNE OTTO, University of Wisconsin

The present study was the first of a series to be devoted to the examination of children's approaches to the task of conceptualizing the main idea in reading. Because existing studies present neither concise methodological guidelines nor definitive descriptive data regarding childrens' performance with a main idea task, it was necessary to focus simultaneously upon the development of an operational approach—e.g., operational definition of the main idea, appropriate reading materials, directions to subjects, a scale for evaluating responses—and the collection of descriptive data. The decision, then, was to devote this initial study to the description and qualitative analysis of children's responses when they are required to formulate their own main idea for a selection.

Three basal paragraphs were developed, each of which embodied an unequivocal but unstated main idea which was operationally defined as a combination of two elements: a general topic—e.g., "birds build nests," and a restriction derived from specific details—e.g., "... in different places." Each paragraph was then rewritten at six different levels, ranging from basal to sixth grade level according to readability formulas. The subjects were 288 children from grades one through six, of whom half read base level and half read grade level materials. Pilot studies revealed that the directions to simply "tell what the paragraph says" were most useful. Responses were evaluated by three judges with the aid of a 7-point scale that was devised to quantify main idea responses ranging from irrelevant to adequate. An individual's score was the sum of the values assigned to the three paragraphs.

A tally of scaled main idea response values revealed that, in general, the subjects had extreme difficulty stating the main idea of the paragraphs. Analysis of variance showed that ability increased across grades, but the tendency was to do with the basal level materials. Implications for instructional practice and further research are discussed.

11:00 to 11:55 A.M. **SESSION 3.5** **DIVISION B**

EARLY CHILDHOOD EDUCATION

CHAIRMAN:

George A. Beauchamp, Northwestern University

PAPERS:

Cognitive-affective Comparisons between Disadvantaged and Nondisadvantaged Kindergarten Age Children

MILTON C. HILLERY, Racine Public Schools, Wisconsin; RONALD H. LINGREN and ROBERT C. REMSTAD, University of Wisconsin —Milwaukee

The building of an adequate curriculum for primary age disadvantaged children requires a greater fund of knowledge than now exists about the specific cognitive and affective differences existing between disadvantaged (as defined by Head Start standards) and non-disadvantaged children. To identify these differences, a year-long descriptive study of four groups of kindergarten age children (30 in each) was made. These groups represented disadvantaged and non-disadvantaged populations from both a large, metropolitan center and a small town-rural setting.

Based on a theoretical structure developed by Sigel, nine different aspects of cognitive-affective processes such as conceptualizing strategies, perception, language development, number concepts, and self-perception were studied. These processes were measured by selected tests or subtests of instruments such as the Illinois Test of Psycholinguistic Abilities, the Frostig Test of Visual Perception, French's Pictorial Test of Intelligence, and several instruments developed or modified by the authors.

Analysis of the data resulted in comparisons between the several group means and the published norms of the standardized instruments, contrasts of the performance of disadvantaged and non-disadvantaged children in both settings, and in developmental trends noted over the ten months of the study. While the wide scope of the study precludes a discussion of the findings in their entirety, the most significant results are discussed.

◆ ◆ ◆

Children's Art Preferences

SUZANNE NELSON, Southwest Regional Laboratory for Educational Research and Development, and GALE DAVIS, Pierce College, Los Angeles

Illustrations are frequently used in instructional materials for young children to stimulate interest and provide environmental referents for textual material. If children prefer one specific artistic style and one technique over others, it is important to identify the preferred style and technique so that they may be used to increase the motivational value of illustrated instructional materials. The purpose of this study, therefore, was to investigate the art preferences of kindergarten children on the most commonly used dimensions of artistic style and technique.

Three stimulus subjects (a house, a bear, and a girl) were prepared for use in the study. Each subject was hand drawn by the first author so that it varied systematically on three dimensions of style (real, cartoon, or stylized) and three dimensions of artistic technique (color, black & white, or line drawing). Each style was represented in all 3 techniques, and each technique was represented in all 3 styles. Thus, 9 illustrations of each stimulus subject were prepared. From among the 9 illustrations for each stimulus subject, 47 five-year old children chose in order the 3 pictures they liked best and those they liked least.

In order to obtain an index of the extent to which children agree in their preferences for the stimulus materials, the investigators computed the coefficient of concordance (w) on the rankings obtained. The average intercorrelation between children (r) was also computed.

The findings clearly indicate that kindergarten children do have significantly consistent preferences along both artistic dimensions: style and technique. Furthermore, these preferences remain consistent across content areas.

The results obtained will be discussed in terms of their implications for the development of instructional materials for kindergarten children.

◆ ◆ ◆

A Behavioral Analysis of Problem-Solving Skills for Young Children

ROBERT J. BERGER and EVAN R. KEISLAR, Southwest Regional Laboratory for Educational Research and Development

The design of materials and procedures for developing the problem-solving capability of young children is one of the objectives of the Southwest Regional Laboratory for Educational Research and Development. Any process or strategy to be taught must be one which is essential to the terminal behaviors identified as "problem-solving." In the formulation of objectives, primary consideration was given to behavior which would likely have generalizable utility in other aspects of the elementary curriculum, as opposed to solving puzzles and games which have little or no apparent transfer to other tasks.

Behaviors identified as meeting the criterion were the classification, ordering, or transformation of objects or events according to a specified set of stimulus dimensions. These behaviors possess a logical relationship to problems in communication, mathematics, science, or any other area of the elementary curriculum. An analysis of classification, ordering, and transformation revealed the necessity of first teaching the concepts associated with the stimulus dimensions. That is to say, one cannot classify by color, shape, size, spatial, relationship, etc. without first having attained concepts such as red, triangle, biggest, above, etc.

A criterion test of childrens' ability to handle these concepts was developed and given to a sample of kindergarten and first-grade children. In addition, instructional sequences were prepared and used with a sample of kindergarten children. The results will be presented and discussed in terms of children's ability to deal successfully with various dimensions of problem-solving, and the formative evaluation of instructional materials and objectives.

◆ ◆ ◆

The Effect of Delayed Admission on Reading Achievement

IRVING H. BALOW, University of California, Riverside

A regularly recurring question in elementary education is the influence of chronological age on reading achievement and the efficacy of delaying admission to first grade for children who appear less mature than the norm. In this study, 131 children whose parents accepted counsel to delay first grade entrance for one year were matched at entrance to first grade with normal age children. Matching was on sex, IQ, school attended, and a twelve month difference in age. There were 81 boys and 50 girls in each group, mean IQ was 96, and mean age was 84.7 months and 72.7 months. Annual comparisons were made of the differences in reading development between the two groups.

Important findings include (1) the older group scored significantly higher on the Metropolitan Readiness Test at the beginning of first grade; (2) in the second, third, and fourth grades the normal age group scored significantly higher on reading comprehension tests; and (3) the mean standard score of the older group decreased each year while the mean standard score of the younger group increased.

11:00 to 11:55 A.M. **SESSION 3.6** **DIVISION B**

SCIENCE AND MATHEMATICS

CHAIRMAN:

J. W. George Ivany, Teachers College, Columbia University

PAPERS:

Science and Mathematics Instruction in Kindergarten and First Grades: Outcomes in Logical Thinking in Second Grade

MILLIE ALMY, Teachers College, Columbia University

This is a preliminary report of a longitudinal study comparing classes participating in the Science Curriculum Improvement Study of AAAS science program and Greater Cleveland Mathematics programs in kindergarten and first grade, with classes having only the mathematics program in kindergarten, classes beginning both science and mathematics programs in first grade, classes beginning mathematics but not in first grade, and classes who had no special program in either mathematics or science in either kindergarten or first grade.

The study has involved 770 children; 518 beginning in kindergarten in the fall of 1965, and an additional 252 beginning in first grade in the fall of 1966. Individually administered pre-test measures of logical thinking, derived from Piaget, included conservation of number and class-inclusion tasks. The post-test measures supplement these with seriation, matrix, transitivity, and conservation of weight tasks. The Peabody Vocabulary test serves as a measure of verbal ability. Measures of other variables are derived from content analyses of the instructional program, observations in the classrooms and teacher questionnaires.

The present report deals with the patterns of progress in logical thinking in the various instructional programs in relationship to the children's initial status in both the logical thinking and verbal ability measures.

◆ ◆ ◆

Computer Simulated Laboratory Exercises in Qualitative Chemical Analysis

T. T. HOLLEN, Jr., C. VICTOR BUNDERSON, and J. L. DUNHAM, The University of Texas at Austin

A computer-assisted instruction course simulating traditional qualitative chemical analysis exercises has been developed for freshman chemistry courses. Analysis of traditional laboratory exercises indicated that, apart from the development of manipulative skills and techniques, the expected terminal behaviors are the recall and inter-relationship of knowledge as to analytical actions, substances involved, and observable phenomena pertinent to the analysis scheme. Justification for the development of the course was based upon premises that the non-manipulative terminal behaviors of traditional exercises could be produced as well by the simulated exercises and at a considerable saving in student time.

The course, written in Coursewriter I language for use on the IBM 1440 System, is organized into segments or modules corresponding to various ion groups. The program permits the instruction to enter unique unknowns for each student, after which the student is led in socratic fashion toward a conclusion as to the presence or absence of ions. The course may be characterized as intrinsically programmed and response centered in its educational approach. Use is made of constructed responses, discriminatory cues, and reinforcing practices. No provision is made for individual differences through "multiple track" techniques.

A comparative evaluation lends support to the contention of equivalent terminal behaviors and less time for the simulated exercise. Subjects, who were members of three laboratory sections in freshman chemistry classes, were randomly assigned to experimental and control groups. There was no significant difference in terminal behaviors between the groups. The mean time required for completion by subjects performing the simulated exercise at the computer terminal was approximately one-half that required by subjects performing the traditional exercise at the laboratory bench.

◆ ◆ ◆

Some Pupil Achievement Effects of Three Modern Mathematics Programs

GERALD L. ERICKSEN, St. Olaf College and Minnesota National Laboratory; and **JAMES J. RYAN**, Minnesota National Laboratory

The general objective was to determine the effects on mathematics achievement of each of three modern programs in mathematics developed for grades 9, 10 and 11. These were the Ball State, UICSM and SMSG programs.

The study included 39 pairs of classes in schools distributed over a five state area. Each pair was taught by the same teacher, one with a traditional program, the other with one of the three modern programs. A randomization test, to determine within pair comparability on initial pupil mathematics ability, was used to select these pairs from a larger participant sample. Two post achievement measures were used. One was a test developed to assess general concepts, principles, and relationships common to both traditional and modern programs at each grade level. The other was the mathematics section of the STEP.

For each program at each grade level, comparisons between modern and conventional class pupils were made using a three-way (program, teacher and pupil ability level) analysis of variance and a two-way (teacher by program) analysis of covariance with initial pupil ability as the covariate. Teachers were considered as a random effect in both instances.

Both the ANOVA and ANACOVA indicated significant achievement differences on one of the tests for the tenth grade SMSG ($p < .07$) and the ninth grade Ball State ($p < .01$) and SMSG ($p < .05$) program comparisons—the latter requiring some qualifications due to nonhomogeneity of regression slopes. All differences favored the modern programs. A difference ($p < .01$) favoring higher ability eleventh grade SMSG pupils was also observed, as were teacher by program interactions for the UICSM and SMSG ninth grade programs.

The discussion will consider the implications of the differences in terms of achievement test variance accounted for, the nature of the teacher by program interactions, the variation in results between tests and grades, and the similarity with results obtained for the evaluation of these programs in other years within this project.

◆ ◆ ◆

Age of Entrance Into the First Grade as Related to Arithmetic Achievement

JOSEPH ILIKA, Lamar State College of Technology, Beaumont, Texas

This investigation was designed to assess the influence of age of entrance to the first grade on subsequent arithmetic achievement. Developmental theory holds that an early start will not result in significant gains of long-term duration. This proposition was tested by comparing the achievement of early and late entrants to the first grade not only by grade, but also at age as they advanced through school. Early entrants have the advantage of an extended period of schooling when comparisons are made at age.

Early and late entrants were matched according to sex, intelligence, and social class, with the result that 41 pairs of boys and 49 pairs of girls were available for study. The difference between the age of entrance of the early and late starters averaged about nine months.

The *t*-test was applied to test the significance of the differences between the mean achievements of the early and late entrants in each grade from one through six.

The comparisons at age were made by estimating the mean achievement ages of the early entrants at the same chronological ages as those of the late entrants. The differences between the mean achievement ages of the late entrants and the estimated means of the early entrants were tested for significance by the *z*-test.

All but one of the sixteen mean arithmetic age comparisons by grade favored the late entrants. Differences based on comparisons of boys were significant in the second, third, and fifth grades, while those based on comparisons of girls were significant in the fourth and fifth grades. Differences in favor of early entrant boys were significant at ages 88, 101, and 125 months; differences in favor of early entrant girls were significant at ages 91, 101, 113, and 137 months. These findings support the teaching of arithmetic at an earlier age.

11:00 to 11:55 A.M. **SESSION 3.7** **DIVISION B**

SOCIAL STUDIES

CHAIRMAN:

G. Wesley Sowards, Stanford University

PAPERS:

The Relationship Between Teacher's Dogmatism and Teachers' Concept of the Discipline of History

EMILY S. GIRAULT, University of Michigan

Assumptions underlying this study:

1. That implicit among the various social studies curricula there are contrasting concepts of the nature of history as a discipline.

2. That every presentation of historical material (whether by teacher, curriculum expert, or textbook writer) represents *some* view of the nature of history.

3. That the teacher's conceptual pattern concerning history will affect his presentation of history, his teaching method, his selected learning activities and his course outlines.

Social studies teachers who are high in Dogmatism will differ from those low in Dogmatism in their view of the discipline of history and of the role of the historian. Social studies teachers high in Dogmatism will view history as a body of concrete, objective, unchanging "facts" about the past, and the role of the historian as that of a recorder-reporter of these facts. Social studies teachers low in Dogmatism will see history as knowledge subject to continuing interpretation and revision, and the role of the historian as that of interpreter.

Subjects of this study were participants in thirteen separate NDEA summer institutes at thirteen universities for history teachers.

An instrument was designed and pre-tested by the researcher for the purpose of ascertaining the subjects' view of history and the historian along the dimensions described in the hypothesis. Roughly speaking, subjects' views of history were summarized in such a way as to allow their placement on a continuum between these extremities:

—History is a finished product. It is the story of the past which never changes.

and . . .

—The past is never fixed. Every society rewrites its history as often as the society changes.

Teachers registered their responses to this instrument as well as to the Rokeach Dogmatism Scale.

Analysis of the results indicates a correlation between dogmatic thinking in the teacher and the teacher's view of history. The data confirm the hypothesis.

Personality variables of the history teacher may be a determining factor in the design of instructional objectives and in their facilitation. Since teachers high in Dogmatism and teachers low in Dogmatism have contrasting views of history, the course content and instructional objectives for classes in history by these groups of teachers may also differ in similar manner.

◆ ◆ ◆

A Study of Estimated True Achievement Gains in Eighth Grade Social Studies, Atlanta, Georgia, School Systems, 1966-67

OWEN SCOTT, The University of Georgia

The purpose of the study was to ascertain estimated true gains in achievement over each of two semesters in a newly introduced eighth grade social studies course, "Changing Culture". Reading achievement data and pretest-posttest achievement data over each semester were obtained for a purposive sample of 311 pupils typical of eighth grade pupils in the school system. On the basis of performance on the Metropolitan Reading Test, Advanced, given in September, 1966, each pupil in the sample was classified as a superior, average, or below average reader. With the sample divided into three groups on the basis of pupil reading level, analyses of estimated true achievement gains were made. The information obtained pertained both to group and individual achievement gains.

Results showed that each group of readers improved significantly in achievement each semester, that the amount of gain varied with the reading level of the group, and that each group achieved at a higher level at the beginning of the semester than lower groups at the end. Also, as we expected, the number of individuals in a group making significant achievement gains varied with the reading level of the group, with almost all superior readers making significant achievement gains and few below average readers doing so. The analyses also yielded suggestive information concerning portions of the course learned most and least effectively by each group.

The somewhat unique features of the study include the use of reading achievement data in the analysis of achievement gains and consideration of regression and score unreliability influences in the quantitative estimates of such gains. Currently, data obtained from the study are being used by instructional supervisors and curriculum workers in revising text materials and instructional methods in the course.

A Comparison of Inductive and Deductive Materials for Teaching Economic Concepts to Culturally Disadvantaged Fourth Grade Students

B. J. DOOLEY, University of Georgia

The purpose of the study was to discover whether the inductive and deductive materials used by regular elementary school teachers to teach economic concepts would result in differences in culturally disadvantaged fourth grade students' learning. The two new curricula materials were authored by William Rader at the University of Chicago and Lawrence Senesh at Purdue University.

A pilot study was conducted to determine the reliability of a 90 item multiple choice instrument. The results of this pilot provided data for revision of the evaluation instrument. A 55 item evaluation instrument was used in pre- and post-tests.

A total of 18 fourth grade classes were randomly assigned to one of the two treatment groups. Each of 484 subjects was administered a pre- and post-test of economic concepts. A 75 item college level economics test was administered to each teacher to assess economic knowledge.

All instruments were statistically analyzed with the 90 item pilot test having a Kuder-Richardson reliability coefficient of .67, the 55 item program economic test having a pre- and post-test reliability coefficient of .83 and .89, respectively. The 75 item economics test administered to the teachers had a reliability coefficient of .85.

A general linear regression model of least squares analysis was utilized in variance analysis of main and interaction effects. Treatment, race, and sex main effects were significant at the .01 level ($P < .01$). First order interactions of treatment × race, treatment × sex, and race × sex were significant at the .05 level ($P < .05$).

The t-test was utilized to test significance between mean gains. Negro subjects' mean gains, using the inductive material, were significant at the .05 level ($P < .05$) where compared to white subjects' mean gains using the deductive materials. The inductive method was consistently more effective than the deductive method with all culturally disadvantaged subjects in the study.

◆ ◆ ◆

A Study of Sixth-graders' Comprehension of Specially-prepared Materials on Broad Social Conflicts

JAMES W. MERRITT, Northern Illinois University

Five stories with an accompanying test were used for assessing sixth-graders' comprehension of broad social conflicts. Content consisted of fictionalized accounts of controversies over antivivisection, oleomargarine legislation, musicians' rights, Negro segregation, and mining safety. The test consisted of 125 multiple-choice items which emphasized "higher mental processes." Comprehension was sub-divided: *viz.*, Knowledge of (1) concepts, (2) opinions, (3) reasons and motives, (4) rights and duties, and (5) analogous situations.

Focus was on the sixth grade because expert opinion indicated strong readiness at this level.

The sample consisted of 481 sixth-graders, 184 fifth-graders, 169 seventh-graders. Lower-class children from a medium-size city were included.

Reliability was estimated in a pilot study at .98 (Spearman-Brown). Validity was indicated by correlating interview scores which yielded a Pearson product-moment coefficient of .78 (standard error, .10).

The stories and sub-tests were presented in a standard manner by the same researcher throughout in a limited teaching context which included preliminary motivational discussion for each story, with evaluatory discussion and score-reporting following each sub-test.

Mean scores indicate a sharp rise in ability to comprehend from grades five to six with a gradual rise from grades six to seven to nine. Seventy-nine percent of the sixth-graders achieved a score of fifty percent or better; fifty percent is defended as a reasonable criterion for comprehension.

Test performance correlated fairly high (.71, standard error, .03) with reading scores, somewhat lower (.55, standard error, .04) with intelligence scores, and lower still (.27, standard error, .04) with Sims scores for socio-economic level. The latter measure suggests that the materials are appropriate for widely differing social classes; this is supported by the high interest expressed verbally by the children. Sixth-graders chose fewer prejudiced stereotypes in the false options than did ninth-graders, suggesting more open-mindedness of the former.

The conclusion is that the general use of this type of material in sixth grades can be justified.

12:00 M to 1:55 P.M. SESSION 4.4 DIVISION D

RESEARCH ON CLASSROOM BEHAVIOR

CHAIRMAN:

W. Robert Dixon, University of Michigan

PAPERS:

A Computer Investigation of Characteristics of Teacher Lecturing Behavior

JACK H. HILLER and GERALD A. FISHER, University of Connecticut

N. L. Gage has clearly stated the importance of studying teacher performance in its natural setting. The difficulties in generalizing from laboratory studies of rat learning, as well as the unsuitability of some variables and the absence of others, present serious problems. Since explaining by lecture is a teacher function, data have been collected for its study.

A representative sample of ten lectures from a collection of about fifty were made available by the Stanford Center for Research and Development in Teaching. Each of the ten teachers had lectured to his regular twelfth grade class for fifteen minutes in an effort to explain contemporary historical information presented in a magazine article read by the teachers before class. The students were given a standard multiple choice test after the lecture to provide a criterion measure of teacher effectiveness. Test scores for a single lecture presented to all classes by tape recorder were used as a control for class interest and ability.

Three categories of variables were theorized to represent important characteristics of explanatory lecturing:

1. Speech comprehensibility (e.g. sentence length, complexity);
2. Motivational;
3. Structural (e.g. causation, effects, and principles cued).

A computer program was written to measure the frequency of occurrence of words in the lectures believed to represent variables within the three general categories. Many of the variables yielded significant correlations with the criterion, and a trend of relationships was apparent as well. The overall analysis is that the poor lecturers presented too much information (for example, total number of words $r = -.67$, p less than .05 under a two tail test); they drew attention to too many things ($r = -.76$); and projected vagueness through an overuse of qualifiers ($r = -.88$).

Analysis of Interrupted Time-Series Experiments

GENE V. GLASS, University of Colorado, and THOMAS O. MAGUIRE, University of Alberta

The purpose of the paper is to present a class of inferential statistical models appropriate to the analysis of interrupted time-series experiments (Campbell & Stanley, 1963). These models provide inferential statistical techniques for testing hypotheses concerning either the shift in level of a time-series or in the direction of its drift associated with the introduction of an experimental manipulation. The simplest model of the class is due to G.E.P. Box and George Tiao (1965). This model was extended to cover a broad range of applications by Dr. Tiao serving as a consultant to the project being reported upon in the paper. (The authors have developed a series of computer programs to perform the analyses and have evaluated the model on data from experiments in education, psychology, sociology and medicine.)

Illustrative analyses are performed on data from an experiment on the effects of alteration of deprivation schedules in an animal learning experiment, data concerning the effects of revision of divorce laws on the divorce rate, and data from an experiment on the effects of social reinforcement of a hyperactive child.

The inferential statistical models are discussed in relationship to the design of interrupted time-series experiments. Analyses for multiple-group time-series experiments and dependent-group time-series experiments are indicated.

This new class of inferential statistical techniques is shown to provide a powerful analysis of data which heretofore have been subjected to invalid analyses.

◆ ◆ ◆

A Study in the Development of a Reciprocal Category System for Assessing Teacher-Student Classroom Verbal Interaction

RICHARD L. OBER and SAM WOOD, University of Florida; ART ROBERTS, Rocky Point, New York Public Schools

This research, a pilot study, was an attempt to modify the Flanders System of Interaction Analysis so that it could (1) be used to train high school students to analyze their own verbal behavior and (2) assess more comprehensively the broad spectrum of student classroom verbal behavior as well as teacher verbal behavior. Specifically, the purposes of the research were twofold:

1. To determine if high school students could learn to use a modification of the Flanders System of Interaction Analysis. Special attention was

given to the identification of instructional problems related to this activity.

2. To develop a reciprocal category system of interaction analysis consisting of an equal number of categories for both teacher and student verbal behaviors.

Eight high school English classes were selected to participate in the study. One of the researchers met with each class for five consecutive days during the regularly scheduled English class period. Seven of the classes were taught to use a 13-category modification of the Flanders System of Interaction Analysis while the one remaining class was taught to use a reciprocal category system developed by Ober. Reliability checks were obtained for all subjects of the eight classes studied. Students' subjective comments and reactions regarding the value of these training experiences were also obtained.

The study produced the following results and/or conclusions:

(1) With a minimum of five clock hours, high school students can be taught to use a modification of the Flanders System of Interaction Analysis.

(2) A reciprocal category system has been developed. After only limited use and testing, the system shows promise as a means for assessing more comprehensively the broad spectrum of student classroom verbal behavior. The system presently consists of 19 categories—nine each for teacher and student verbal behaviors and a single category for silence or confusion.

Additional data are currently being collected in order to determine the effectiveness of the reciprocal category system as a means for helping preservice and inservice teachers to control their verbal teaching behavior.

◆ ◆ ◆

Some Effects of Observers' Beliefs on Classroom Observations of Teachers' Behavior

WM. NICHOLAS STOFFEL, Saint Leo College and BOB BURTON BROWN, University of Florida

This study investigated the effects of the personal and educational beliefs of observers on the reliability and validity of their observations of the classroom behavior of teachers.

This study involved 557 observer-subjects drawn from the faculties of six teacher education institutions and their cooperating public schools in California, Florida, Illinois, New York, and Wisconsin. Each of the observer-subjects responded to the *Personal Beliefs Inventory,* the *Teacher Practices Inventory* and the *Dogmatism Scale,* and observed one or more student teachers at work in public school classrooms, recording their observations on the *Teacher Practices Observation Record.* A total of 1507 observations were made of 407 student teachers.

Relationships were found between observers' beliefs scores and the observational scores given by the observers to indicate that reliability and validity are affected in predictable directions.

The results of this study illustrate the importance of knowing, controlling, or accounting for observer bias in interpreting data yielded by so-called "objective" observational systems.

12:00 M to 12:55 P.M. SESSION 4.6 DIVISION A

SPECIAL PROBLEMS IN ADMINISTRATIVE RESEARCH

CHAIRMAN:

Thomas J. Sergiovanni, University of Illinois

PAPERS:

The Social Sciences, Educational Administration and the Sociology of Knowledge

EMIL J. HALLER, Ontario Institute for Studies in Education

This paper reports a pilot project for a sociological study of the "theory movement" in educational administration. It was designed to test the use of citation analysis as a tool for answering questions about the nature of the social science influence on research in educational administration. Three questions were proposed: Which of the social sciences is currently having the greatest impact on administrative research? How recent are the ideas borrowed from the social sciences by administrative researchers? Which social scientists are currently having the greatest impact on the thinking of these rearchers?

A citation analysis of all research published in the Educational Administration Quarterly was performed. By the citation criterion, sociology is clearly the most influential discipline, far ahead of psychology, social psychology, political science, economics, and anthropology, which followed in that order. Material drawn from the social sciences appears to be older than that drawn from education itself. Finally, the most influential sociologists seem to have been Blau, Gouldner, Weber, and Parsons.

Two implications of these findings are discussed. Possible reasons for the dominance of sociology are proposed, and several hypotheses are suggested for the major research for which the study served as a pilot. Second, possible dysfunctional aspects of the theory movement are suggested.

◆ ◆ ◆

Patterns of School Research in Thirty-Three Suburban School Systems

ROBERT M. RIPPEY, The University of Chicago

Demographic and organizational variables were examined in thirty-three school districts in the Chicago suburban area. These variables were related to the amount of research and evaluation practices reported in these schools by means of multiple regression analysis. Over fifty percent of the variance of the research and evaluation activity could be explained in terms of three variables. These were 1) the presence of a research director, 2) the size of the district, and 3) the percentage of rental housing in the district. The relationship between total amount of research and evaluation activity and school size was a negative relationship with the smaller schools i.e., 1500 to 2500 students doing the most research and evaluation. Following the initial survey, individual interviews were conducted with numerous personnel in the seven schools which reported substantially more research and evaluation activity than was predicted in terms of the three factors previously mentioned. The results of these interviews show likenesses and differences among these exemplary schools. The paper will consist essentially of three parts.

1. A summary of the types of research conducted in these districts, following the conceptual structure of Lazarsfeld and Sieber.

2. The results of the analysis of the data attempting to predict the amount of research and evaluation activity on the basis of school and community factors.

3. A summary of the results of the interviews conducted within the exemplary schools.

◆ ◆ ◆

Professionalization Strategies: A Comparative Study of National Teacher Organizations in Canada and the United States

HERMAN A. WALLIN, The University of British Columbia

Historically the NEA has enjoyed a much larger support-base in terms of membership and financial wealth than its rival, the AFT. The national teacher organization in Canada, the CTF, seems to have been an impotent force in Canadian education. In recent years, however, the AFT has made rapid gains over the NEA in achieving support from U.S. urban teachers, while the CTF faces critical unity problems. In what ways have the professional strategies and issues of each of these national teacher organizations changed since 1955? What factors appear to be responsible for the different patterns of national involvement between the teachers of Canada and the U.S.A.?

To secure comparative data both documentary analysis and interview techniques were employed. Documents which were studied included budgets, editorials, agenda of annual meetings, and recruitment literature. Interviews were obtained with officers of all three organizations in Washington, D.C., and Ottawa.

◆ ◆ ◆

Open Door Academy: Adaptation to Insufficient Enrollments by a Non-Public School

JOSEPH M. CRONIN, Harvard Graduate School of Education

For certain types of organizations, such as private schools, a steady flow of clients may or may not be assured. Carlson developed a typology of "wild" and "domesticated" organizations according to whether a client had any control over participation and an organization control over admissions. His particular concern was the response pattern of organizations and clients when neither could initially exercise any control. Burton Clark in *Open Door College* studied a second type of school wherein the clients, by exercising program options, effectively reshaped a junior college. This paper reports a study in which a third type of organization (what Clark calls selective-voluntary or Carlson a Type I "wild" organization) is studied over the the first four years of existence.

The school studied was a private academy which intended to establish a reputation as an elite liberal arts college preparatory school in a growing suburban area. Although the school held to rigorous admission standards at first, a combination of transportation and other problems discouraged students from enrolling after being selected. In this case a "wild" organization adapts by chartering busses, modifying the program, lowering admissions standards, and extending recruitment efforts. But the acceptance of less talented clients forces further program dilution or revision when the issue becomes one of clientele retention.

Among the findings is the observation that clientele retention problems may provoke parallel behavior patterns on the part of school officials in three theoretically discrete organizational settings, mainly because schools cannot ignore financial commitments by placing highest premium on selectivity. The "wild" or selective-voluntary type of school may include several sub-types, new and established, some lean and hungry, and others sleek and complacent with waiting lists and little impetus towards program adaptation. The responses of the case under study yield some insights about the network of constraints which shape a non-public school and influence clientele participation.

12:00 M to 12:55 P.M. **SESSION 4.7** **DIVISION A**

SUPERVISORY BEHAVIOR AND TEACHER MOTIVATION

CHAIRMAN:
Terry L. Eidell, University of Oregon

PAPERS:

Motivating Factors in Supervision
ARTHUR BLUMBERG, Syracuse University

The purpose of this study was to describe and analyze those factors in supervision that affect the motivation of teachers.

The methodology was an adaptation of Herzberg's technique. In this work, depth interviews were conducted with engineers and accountants in large organizations. They were asked to recall particular times, in their work, when something very favorable and something very unfavorable had happened. The assumption was that people would try to select and arrange an environment where the favorable (motivational) factors would be more likely to occur and to avoid an environment where unfavorable factors occurred.

The results indicated two classes of favorable factors: (a) motivational, which tended to be associated with the work itself, and (b) hygienic factors which, in themselves, did not motivate strongly, but provided a basis upon which motivational factors could operate.

In the present study a population of 210 teachers responded to open-ended questionnaires instead of depth interviews. The questions were focussed on supervisory behavior rather than the total work situation. The respondents were also asked to rate the quality of their supervision.

The results were content-analyzed. They supported, in general, the results of Herzberg's work. Eighty per cent of the "favorables" were associated with work-oriented aspects of supervision; twenty per cent with hygienic or social-emotional aspects. Ninety-three per cent of the "unfavorable" were hygienic in nature; seven per cent had a work orientation. Those who rated the quality of their supervision as high placed less emphasis on the receipt of praise as a favorable factor than did those who rated it low. Those who gave a low rating to their supervision indicated more strongly than the highs a lack of trust by and support from their supervisor.

This study, though exploratory, seems to justify further research into the effects of supervisory behavioral styles on teachers and teaching.

erceptions of the Supervisory Behavior of Secondary School Principals

LAWRENCE J. MARQUIT, Educational Research Council of Greater Cleveland

Purpose is to analyze how teachers and principals perceive the extent of principals' supervisory behaviors; and to analyze the relationships that may exist between their perceptions and certain demographic variables.

An original instrument, the Opinion Inventory of Supervision (OIS), was developed and designed to measure teachers' and principals' perceptions of the frequency of each of nine categories of supervisory behavior: curriculum development, instructional organization, staffing, providing instructional aids, orientation of new instructional staff, providing in-service education and professional growth, coordinating special services, school-community relations, and evaluating. Each category was represented by four randomly paced items, comprising Part I of the OIS. Each item required the respondent to make a value judgment of the frequency of the supervisory behavior in question, as: (1) never, or almost never, (2) rarely, (3) sometimes, (4) often, and (5) always, or almost always. Part II of the OIS requested the respondent to indicate his demographic data as to sex, age, experience, educational background, and tenure status in the case of teachers; and sex, age, experience, and prior professional status in the case of principals.

The responses of 696 teachers were compared with the responses of their 15 respective principals to determine the extent of significantly different perceived frequencies and the extent of rank-order correlation.

The responses were also analyzed to determine the extent of significant differences and correlations within each demographic variable between teachers and between principals, and to determine the existence of relationships between principals' demographic variables and their respective teachers' scores.

Results: The data revealed that (a) principals' scores on the OIS were significantly different from (higher than) their respective teachers' scores; and (b) age, experience, tenure-status, and school size were more significantly associated with OIS scores than were educational backgrounds or sex.

◆ ◆ ◆

Supervisory Style of Principals and Teacher Autonomy

JOANNA JENNY SAMUELS, Anoka-Ramsey State Junior College (S. Jay Samuels)

Samuels reported (AERA, 1966) that as school district size increased the autonomy of the elementary classroom teacher to make decisions in the professional domain decreased. To explore further other factors impinging on teacher autonomy, this study investigated the effects of principals' supervisory style on the classroom teachers' scope for decision-making in the classroom.

Data for the study were collected by means of a questionnaire sent to 1,887 teachers representing 224 elementary schools located in 13 school systems ranging in size from an average daily attendance of 2,431 to 83,101. Questions related to teachers' scope for decision-making and teachers' evaluation of their principal's leadership style.

Analysis of indicators of teacher autonomy indicated that a significant relationship exists between the perceived strictness of principals and the professional autonomy of elementary school teachers: the strict principal constrains teachers' scope for decision-making in the classroom.

Further analysis revealed that strict principals are more often found in larger districts. In order to determine whether the decreasing autonomy of the classroom teacher is due to the fact that he works in a large district or is the result of his working under a strict principal who, incidentally, is more often to be found in the larger district, one must ask: what happens to the relationship between district size and professional autonomy when we study the autonomy of teachers working under strict and non-strict principals separately?

Such analysis reveals that constraint of teacher autonomy is more a function of district size than of the supervisory style of the principal.

Regardless of the principal's style of supervision, a strong relationship between large district size and low professional autonomy persists.

◆ ◆ ◆

An Assessment of Teacher Work Values

ROBERT J. COUGHLAN, Northwestern University

This study classified teachers in four high schools according to the strength of their identification with each of the value systems undergirding bureaucracy, professionalism, and the informal work group. The concept of work values was used as a way of describing the dissimilar value-orientations brought into the work situation by the teachers. Theory and previous research on professional work groups in formal organization had suggested as relevant the following clusters of values: 1) Professional—held by teachers who identify themselves with the values and goals of professionalism; 2) Organizational—held by teachers who identify themselves with the values and goals of bureaucracy; and 3) Social—held by teachers who identify themselves with the values and goals of their primary work group.

A 45-item, modified paired-comparison schedule was constructed to measure teacher work values. The instrument was administered to a total of 192 teachers in the four sample schools. The data from this administration were factor analyzed to determine and describe the different types of items in the schedule. The factor analysis yielded five bi-polar dimensions or factors which accounted for 27.6 of total variance among all items. Beta-weighted scores of the teachers on each of the factors were then subjected to a method of pattern or profile analysis (D. R. Saunders' syndrome analysis). The results of this analysis, coupled with an examination of comments written by teachers in response to questions about their achievements and aspirations, provided bases for the identification and naming of five basic types of teachers with respect to work values.

The five types were essentially consistent with the Professional, Organizational, and Social constructs originally posited in the research. However, the findings also sharpened the Professional and Organizational types by indicating that finer distinctions could be made than those explicitly proposed, i.e., two types of Professionals and two types of Organizationals emerged from the data. A sixth group of mixed types of teachers, or Marginals, was also identified.

12:00 M to 12:55 P.M. **SESSION 4.8** **DIVISION C**

CONCEPT LEARNING I

CHAIRMAN:

Joe L. Byers, Michigan State University

PAPERS:

The Effect of the Number and Level of Conceptualization of the Defining Attributes on the Difficulty of a Conjunctive Concept

MICHAEL B. BERZONSKY, University of Toronto

When teaching concepts, teachers find that some concepts are more difficult than others for their students to learn and understand. Therefore, in order for teachers to be most effective, they must have information regarding not only the difficulty of a concept, but also knowledge concerning what factors contribute to the difficulty of a given concept. This study is an effort to determine whether concept difficulty is a function of certain specifiable characteristics of a conjunctive concept.

An analysis of the research on concept difficulty suggests that a conjunctive concept may differ in difficulty on at least two dimensions: (a) number of defining attributes, and (b) the level of conceptualization (LC) of the defining attribute. The LC of a concept's attributes refers to the extent to which an individual has learned the stimulus situations which represent a given attribute.

Two pools of biological conjunctive concepts were assembled. The first group included concepts with three attributes per concept and the second pool consisted of five or six attribute concepts. A consensus of three biologists determined the attributes of each concept.

103 high school subjects (*S*s) were then randomly assigned to one of two groups. Group I of the *S*s was tested on the attributes of the concepts in order to determine the LC of the attributes of each concept. The index of concept difficulty for Group I was the ability to identify *instances embodying attributes of the concepts*. On the basis of the information obtained on Group I the concepts were then categorized in a 2 × 2 factorial design. Group II of the *S*s were then tested on the actual concepts in order to compare the difficulty of concepts having either three attributes or five to six attributes, each subgroup of which had two levels of conceptualization. The index of concept difficulty for Group II was the ability to identify *instances of the actual concepts.*

Four scores were then computed for each S in Group II based on the four categories of concepts, classified in terms of the information obtained on Group I concerning the number and LC of the concepts' attributes. An analysis of variance procedure was then applied to scores of the *S*s in Group II.

The results of the analysis indicate the following: (a) concept difficulty was inversely proportional to the level of conceptualization of the concepts' attributes ($p < .01$); (b) concept difficulty was not differentially affected by the number of defining attributes in each concept; and (c) the difficulty of a conjunctive concept was a function of the interaction ($p < .01$) of the number of defining attributes and the LC of the concepts' attributes.

◆ ◆ ◆

Concept Learning as a Function of Type of Material and Type of Classification

JAMES G. RAMSAY, New York University and HERBERT J. KLAUSMEIER, University of Wisconsin

Sorting of verbal instances on the basis of shared association values was compared with sorting figural instances on the basis of shared values of figural dimensions. The figural instances were 16 patterns derived from four bi-valued dimensions. The verbal instances were 16 nouns which had been shown to be associated with four adjectives of sensory description. Two hypotheses were tested: (1) that difficulty in correctly sorting instances would be a direct function of complexity of the classification; (2) that figural instances would be more difficult to categorize correctly than verbal instances. Three classifications (termed I, II, and III) of increasing complexity were used; complexity was defined as the length of symbolic rule needed to describe the classification.

The *Ss*, 36 male college students, were run individually in one of the six treatments with six *Ss* randomly assigned to each group. The experiment was a 2×3 factorial design with two types of material (figural or verbal) and three classifications of the instances (I, II, and III). The task was to learn to categorize correctly 16 instances into four groups.

An analysis of correct responses revealed that type of classification, type of material, and the interaction of these effects were each significant sources of variation ($p < .001$). Further tests showed the order of difficulty of the classifications (from most to least difficult) was $III = II > I$ for the figural material and $III > II = I$ for the verbal material. For classifications III and II, figural material was more difficult than verbal material, while for I there was no difference between types of material. The hypotheses were supported, but needed to be qualified because of the significant interaction. The results allow findings of concept learning studies with figural material to be meaningfully related to studies using verbal material.

◆ ◆ ◆

The Effect of Problem Complexity upon Dimension Selection and Associative Learning in Concept Attainment

J. DOUGLAS OVERSTREET and J. L. DUNHAM, The University of Texas at Austin

Richardson and Bergum suggest that multiple category concept learning tasks may be divided into three sequential steps: (1) identification of dimensions and their values; (2) identification of relevant dimensions; and (3) the association of the relevant category name with the defining characteristics of the category. They found that the associative learning stage accounted for the overwhelming majority of total trials to criterion. Since Richardson and Bergum used a nine-category conjunctive problem with two relevant and one irrelevant three-valued dimensions, it would seem that their problem was such as to maximize stage 3 difficulty while minimizing the difficulty of stage 2.

The present study attempted to extend and refine the Richardson and Bergum analysis by independently varying (a) the absolute number of associations to be learned, and (b) the number of possible pairs of dimensions from which *S* must select the one pair which is relevant. This was accomplished by using problems in which there were either two or three values per dimension, and either one or two irrelevant dimensions. The general expectation was that the two independent variables would have differential effects on the dependent measure, trails to criterion, in different stages. Specific hypotheses were: (a) an increase in the number of irrelevant dimensions should increase problem difficulty in stage 2, while (b) an increase in the number of values should increase problem difficulty in stage 3.

In the procedure commonly used in concept learning studies, *S* is instructed concerning the dimensions and their values, thereby eliminating stage 1. Stimuli are then presented in a random order. Under these conditions *S* must retain both the characteristics of the stimuli and their correct categorization. In Richardson and Bergum's procedure, and in this study, additional constraints were placed upon stimulus presentation: stimuli were presented randomly within a given "order" where an order is defined as one instance of each concept. This method allows presentation of stimuli without their associated response, thus avoiding contamination of stages 2 and 3.

The results of analyses of variance of the trials to criterion measure in each stage of problem solution supported the hypotheses. These results indicate that task variables can be shown to effect differentially the difficulty of steps necessary for successful concept attainment.

◆ ◆ ◆

Testing a Concept Identifiation Theory Appropriate to Variable Stimulus Sequences and a Theory for a Constant Sequence

JOHN W. COTTON, University of California, Santa Barbara

100 Howard University students were randomly assigned to two groups of 50. All subjects received 24 trials on a three-dimensional binary concept task in which the relevant dimension was whether letters appeared upright or sideways on a projection screen. Irrelevant dimensions were the

alphabetic character used ("a" or "b") and the case (upper or lower) in which they were printed.

The Variable Group received a different random sequence of stimuli for each subject, subject to the restriction that no single stimulus, such as α, appear more than six times out of the 24 trials. The Constant Group received the same sequence for every subject. In each group the correct answer ("0" or "1") was shown after each trial. All subjects were run individually, making oral responses during the 9 sec. period that a stimulus slide was presented. Before and after the 24 trials, each subject responded to a slide containing all eight possible stimuli but received no knowledge of results.

Bower and Trabasso's concept identification model was modified to take into account the *congruence* of successive stimuli. (If the relevant dimension and h irrelevant dimensions are the same in value from one trial to the next, the number of congruent dimensions is called: $i = h+1$. The corresponding statement holds for changed dimension values.) The Bower-Trabasso transition matrix T then was replaced by three matrices, T_i, for the three possible i values. Thus the Constant Group prediction for a given trial depends upon a specific arrangement of different T_i's rather than a power of T, as for the Random Group, to which the Bower and Trabasso model applies directly. Parameter estimation procedures supported each model for the cases studied, and much more marked early trial fluctuation in per cent correct occurred for the Constant Group than for the Variable Group.

1:00 to 2:55 P.M. **SESSION 5.3** **DIVISION C**

COOPERATING ORGANIZATION: INTERNATIONAL READING ASSOCIATION

BEGINNING READING

CHAIRMAN:

Helen K. Smith, University of Chicago

PAPERS:

Cognitive Factors Associated with Achievement in Reading at the Elementary Grades.

JAMES C. REED, Indiana University Medical Center

This study investigated the relation of Verbal tests of the WISC, Performance tests of the WISC, and selected tests of the Halstead-Reitan Neuropsychological tests to reading achievement at Grade I and Grade III.

Children in three public schools were administered the WISC and the Halstead-Reitan Neuropsychological tests during their first semester at Grade I. At the end of Grade I, they were given the Gates Primary Reading Achievement Test. At the end of Grade III, the same children were retested with the Gates Advanced Primary Reading Achievement Test.

The data analysis was as follows: For each school a regression equation was calculated, with Grade I reading score as the dependent variable for the Verbal tests of the WISC, the Performance tests of the WISC, and the Halstead-Reitan Neuropsychological tests. The same procedure was followed using Grade III reading score as the dependent variable. Comparisons were made among the regression equations for each group of tests and for each school. Within each group of tests comparisons were made to determine whether the test having highest weight in the regression equation as the same for each school.

Differences in the proportion of variance due to regression were found among the three sets of tests. However, these differences were dependent upon the particular school, as well as whether reading score Grade I or reading score Grade III was the dependent variable. The Verbal tests of the WISC gave a better prediction of reading at Grade III than at Grade I. For the Halstead-Reitan tests and the WISC Verbal tests, the single test making the largest contribution to the regression equation varied with school. Of the WISC Performance tests, Picture Arrangement had the highest standardized regression coefficient at each school.

The magnitude of the relation of cognitive factors to reading depends upon the school. Within these tests, a single subset of skills which are associated with reading cannot be identified for different samples of school children.

The Effect of Direct Instruction in Listening on the Listening and Reading Comprehension of First Grade Children.

ELIZABETH A. THORN, North Bay Teachers' College, Ontario Department of Education.

The purposes of the study were to assess the extent to which the listening comprehension of first grade children was improved through participating in a program of direct instruction in selected listening skills, and to assess the extent to which such instruction affected achievement in reading comprehension.

The listening and reading achievement of three groups of first grade pupils was compared. A program of instruction in listening was developed, and systematic lessons were taught to the experimental group during the first grade year. During the same period, one of the control groups followed the usual language program, while the other was provided with a special program to determine the Hawthorne effect. Neither control group received direct instruction in listening. Achievement in listening was measured at the beginning of the school year and again at the end. The three groups followed the same reading program. Reading achievement was measured at the end of the school year. Using analysis of covariance, change in the listening and reading achievement in the experimental group was compared to that in the control groups.

A difference in listening achievement significant at the .003 level of confidence was found between the experimental and control groups. No significant differences were found between the control groups. A difference significant at the .06 level of confidence was found between the reading achievement of the experimental and control groups. No significant differences were found (P-.96) between the control groups.

The findings appeared to substantiate the following conclusions: The listening comprehensions of first grade pupils of the type used in this study are significantly improved by direct instruction in listening.

Instruction in listening, under the conditions of this study, has a positive effect on reading achievement, although it does not produce highly significant improvement in reading comprehension. The pattern of instruction used in basal reading programs, when adapted for use with aurally presented material, improves listening comprehension at the first grade level.

◆ ◆ ◆

Children's Reactions to Illustrations Portraying Characters in Integrated Situations

SAMUEL WEINTRAUB, University of Chicago

The purposes of this study were: (1) to discover whether children noticed differences between two sets of illustrations of the same story, one set portraying an integrated situation and the other picturing all Caucasians; and (2) to note and compare children's reactions to various characters in multiethnic and in all-Caucasian illustrations.

A total of 88 children, 44 Negro and 44 Caucasian, were individually interviewed, and responses to three selected sets of illustrations recorded and analyzed. Subjects were randomly selected from a kindergarten, second, and fourth grade in each of three different schools. One school served an all-Negro, lower socio-economic class neighborhood; the second was located in an integrated middle-class area; and the third drew pupils from an all-white upper middle-class section.

Pupils were presented two sets of illustrations on each of three different situations: play, school, and family. In the play and school situations, one set of pictures showed characters in an integrated setting, while the other portrayed the same situation with all-Caucasian persons. In the family situation, one set of illustrations was of a Caucasian family and the other, a Negro family. Paired sets of pictures were presented in alternate order to subjects.

The structured interview format called for subjects to respond to three different sets of questions for each paired set of illustrations: (1) questions dealing with the story and characters in one set of pictures; (2) a similar series of questions on its paired set; and (3) questions calling for a comparison of the two sets of illustrations.

Data were analyzed by school, sex, race, and grade level. Among the results were findings suggesting that younger children and boys tended to identify with characters of the same sex, but older Negro girls appeared to identify with racial characteristics.

◆ ◆ ◆

Visual and Auditory Modalities Related to Two Methods for Beginning Reading

HELEN M. ROBINSON, University of Chicago

The purpose of this study was to determine the relative reading progress made by pupils differing in visual and auditory aptitudes when they were taught by two approaches to reading.

The subjects were high and low visual and auditory perceivers chosen from 232 pupils entering first grade in two school systems using the basal reader approach, and 216 subjects from two other systems using the Hay-Wingo approach, which is strongly phonetic. Instruction was given by 22 teachers in 11 schools in four districts.

The tests for visual discrimination were three shown by Goins to place on a factor, P_2, including reading achievement. The Wepman auditory discrimination was given to high and low visual perceivers. Children failing visual and auditory screening tests were eliminated. The criterion tests were the Metropolitan Achievement Test—reading sections—and the Gray Oral Reading Test. At the end of grade 3, the Metropolitan and Huelsman Word Discrimination Tests were given.

The scores were analyzed using multivariate and univariate analysis. As hypothesized, there was no statistically significant difference between high visual-high auditory pupils taught by predominantly sight and phonic approaches. Likewise, there were no significant differences between the two low groups. Contrary to the hypothesis, the high visual-low auditory group made no greater progress by the sight than by the phonic approach. The low visual-high auditory group made greater progress ($P=.06$) in silent reading at grade 1 by the phonic method, but only a tendency ($P=.10$) remained at grade 3.

Conclusions are based on ancillary data secured from analyses of the visual, auditory, and intelligence tests, of the total population as they help to explain the findings for the four groups.

1:00 to 1:55 P.M. **SESSION 5.8** **DIVISION C**

VERBAL BEHAVIOR II

CHAIRMAN:

Joanna P. Williams, University of Pennsylvania

PAPERS:

Relations Between Perceptual and Syntactic Control in Oral Reading

LAUREN B. RESNICK, University of Pittsburgh

This study was designed to answer two related questions: 1) Is reading in longer perceptual units necessarily associated with reading in syntactically appropriate units? 2) What is the effect of perceptual strain on the length and syntactic character of reading units?

Each subject read aloud a text projected on a screen. When he reached a predetermined word in the text, the screen was blacked out. The subject continued to "read" beyond the blackout point as far as he could. Each subject read 54 passages. This procedure yielded two scores for each subject: 1) the average number of words read after the light-out position (average Eye-Voice Span or EVS); 2) the number of trials on which he stopped reading at a phrase boundary. There were four groups of subjects: third graders, fifth graders, college students under standard presentation, and college students in a condition of perceptual strain created by projecting the slide upside down.

Under standard presentation, both average EVS and the number of stops at phrase boundaries increased from third grade to college. College students under perceptual strain, however, had mean scores on both measures approximately like third graders. In both college groups, subjects with long average EVS stopped more frequently at phrase boundaries than those with short average EVS. For third and fifth graders there was no significant difference. Pearsonian correlation between average EVS and stops at phrase boundaries was significant only for the college strain group.

While both perceptual and syntactic control in reading increase with age, the relation between the two scores within each group of subjects suggests that only in skilled readers is appropriate syntactic behavior a direct function of extended perceptual span. For skilled readers, a condition of perceptual strain will lower both length of reading unit and syntactically appropriate unitizing. However, perceptual strain appears actually to heighten the association between length and appropriateness of the reading unit for this group. These findings imply that perceptual difficulty alone cannot account for the lack of association in younger readers, and that merely increasing perceptual skill in children would not automatically improve their syntactic behavior.

The Role of Certain Grammatical Relations in the Understanding of Sentences

PANDELIS G. HALAMANDARIS, Indiana University (Malcolm Fleming)

On the basis of the grammatical theory developed by Noam Chomsky, it is reasonable to presume that the different parts of a sentence may not all be understood with equal facility, i.e., with equal speed. One purpose of this study was to determine whether some of the grammatical relations within a sentence were understood more readily than others. A second purpose was to provide systematic, experimentally-based information concerning guidelines for preparing drill materials of the kinds used in second-language learning. Such information supplements other psycholinguistic studies (e.g., Blumenthal, 1964; Gough, 1965, 1966; Miller, 1962; Slobin, 1963, 1966).

Declarative sentences of varying grammatical form were presented to *Ss* who were asked to verify them one at a time by comparing each sentence with the content of a picture that was shown contiguously following the sentence. Speed of verification was taken as an index of speed of understanding. On a given occasion, the sentence was either active or passive, affirmative or negative, and confirmed or falsified by a picture. Further, on those occasions when the sentence was falsified it was falsified either with respect to the subject-of the sentence, the main-verb-of the sentence, or the object-of the sentence. Each subject responded on all of the 192 occasions that were arranged. The experimental design may be described as a "2 x 2 x 2 x 3 x subjects" factorial design.

The results supported those obtained by Miller, (1962), Gough, (1965, 1966), and MacMahon, (1963). Active sentences were verified faster than passive, affirmative faster than negative, in keeping with predictions based on Chomsky's theory.

Further, the results showed that Chomsky's theory was confirmed with respect to facility of understanding as a function of grammatical relation in the case of negative sentences but not in the case of affirmative sentences.

Finally, the facility of understanding was found to be greater in the case of affirmative sentences than in the case of negative ones. This result suggests that the semantic as well as the syntactic component is of significance. That is, meaning as well as structure influences understanding.

◆ ◆ ◆

The Effect of Syntactical Violation Upon the Recall of Form Class

DENNIS L. ANDERSON and JOE L. BYERS, Michigan State University

The purpose of the present study was to determine the effect of syntactical violations upon the recall of various grammatical form classes when the stimuli were presented by audio, visual, and audio-visual modes. It was hypothesized that there would be significant differences in the recall of words according to form class with nouns having the highest recall, followed by verbs, adjectives, and adverbs. An interaction was also expected between grammatical form class and type of syntactical violation.

A free recall procedure was used in which subjects were exposed to word lists representing three levels of syntactical violation: sentences, anagrams, and random strings. The word lists were derived from a source of twelve sentences wherein the words of all grammatical classes were equated for frequency of usage according to the Thorndike-Lorge Word Count Dictionary for Teachers. Subjects were randomly assigned to audio, visual, and audio-visual presentations in which each subject was presented the stimulus materials over six trials. Data were analyzed using a complex factorial analysis procedure with repeated measures.

The results tend to support the findings of other recent studies which indicate that difficulty in learning verbal material may be a function of grammatical class and the context in which words are learned. Significant differences were found among the grammatical classes with nouns being most easily recalled, followed by verbs, adjectives, and adverbs. An investigation of the pattern of variability across trials in the interaction of form class with lists and trials showed a decreasing variance in later trials for sentences and anagrams and an increasing variance for random strings. This may suggest that for word lists where syntactical order is naturally present or can be imposed, relationships among grammatical classes may facilitate the recall of all form classes.

◆ ◆ ◆

Recency and Frequency in the Selection of Words in Sentences

JOAN L. PRENTICE, Indiana University

Frequency and recency of experience with individual words were hypothesized to increase the probability of selection of those words from a pool of synonyms in sentence production. Ninety-Six *S*s were randomly assigned to one condition of recency training involving a high frequency (Hi*f*), low frequency (Lo*f*), or control (C) word. The *S*'s task was to explain the meaning of the specified word first as if to a three-year-old, and second as if to a person not a native speaker of English.

The test task was ostensibly a separate experiment. The *S* listened to a taped dialogue of an inane argument, and was instructed to reconstruct that argument, sentence by sentence, in his own words. The dialogue embedded 10 words of a synonym pool exactly once, in three randomly varied positions. The pool included the Hi*f* and Lo*f* words, but not the C word, of recency training. Sitting behind *S*, the *E* tallied every occurrence of the key Hi*f* and Lo*f* words during *S*'s paraphrase. Each *S* was given two trials.

The Hi*f* word was produced reliably more often in paraphrase than the Lo*f* word, $F = 19.66$, $df = 1,90$, $p < .01$. Because paraphrase is, by definition, idiosyncratic phrasing, and assuming the *Ss* engaged in fact in paraphrase, it seems more likely that the process observed was word selection in natural speech, rather than simple recall of key words. If so, word frequency is one of the factors contributing to word selection in sentence production. Whether a speaker's habits are a result of, or contribute to, cultural norms reflecting word frequency, is unknown.

The word experienced in the recency training task was expected to show a relative increase in occurrences in the paraphrase task. However, recency training had no significant effect on the use of words in paraphrase ($F < 1$).

◆ ◆ ◆

Word Associations and the Recognition of Flashed Words

S. JAY SAMUELS, University of Minnesota

If a subject is required to recognize a tachistoscopically presented target word, e.g., snow, it has been demonstrated that speed of recognition is faster if the target word is preceded by an associate, e.g., white. What has not been demonstrated is the effect of different kinds of word-associations on word recognition speed, and this was the purpose of the present study.

Twenty college subjects were used in a repeated-measure 5 x 5 Latin square design in which the words used with each treatment were counterbalanced. Order of treatment presentation was randomized for each subject.

Three experimental and two control conditions were used. Each condition had two target words making a total of ten target words. To insure that differences in recognition speed were due to the effect of the kind of word-associations presented and not individual differences in familiarity with the word-associations or the target words, the subjects were familiarized with all the target words and the word associates used in the experimental conditions, by having them read them aloud from cards before they were presented with a three-channel tachistoscope.

A paradigm of the conditions will be presented.

Mean recognition speed for each treatment was: facilitation—35.00 ms., control_1—40.00 ms., control_2—47.55 ms., Interference—56.65 ms., Neutral —60.35 ms.

The Newman-Keuls test shows which means were significantly different from each other.

One may conclude that different kinds of word-associations may facilitate or interfere with speed of word recognition.

1:00 to 1:55 P.M. **SESSION 5.9** **DIVISION C**

CONSERVATION CONCEPTS I

CHAIRMAN:

Robert E. Grinder, University of Wisconsin

PAPERS:

The Effects of Sequence on Acquisition of the Concept of Conservation

EDWINA MEYER and EGON MERMELSTEIN, Hofstra University

The objective of this study was to ascertain the effects of sequence on acquisition of the concept of conservation. Sigel's training technique, with its emphasis on sequence, was employed for this purpose. It was hypothesized that there would be no difference between the prescribed sequence of multiple labeling, multiple classification, seriation, and reversibility, and three permutations of the sequence.

In order to test these hypotheses a sample of 80 kindergarten children from Long Island, New York, area was selected and randomly assigned to four situations, group A receiving the conditions in Sigel's prescribed order, and groups B, C, and D receiving the conditions in various random orders. As in Sigel's study, immediately prior to the training all of the subjects were pretested on a series of conservation tasks, and from one to two weeks after the training all the subjects were posttested on these same tasks.

Dunnet's test, a multiple comparison procedure for comparing several treatments (groups B, C, and D) with a control (group A), was employed to directly test the hypotheses. The results indicate that there were no significant differences at the .05 level between groups B, C, and D, and group A on conservation tasks. That is, no one randomly ordered sequence was more effective in inducing conservation than the prescribed sequence.

It is concluded that sequence is not the facilitating factor in inducing conservation.

◆ ◆ ◆

Conservation Training Techniques and their Effects on Different Populations

EGON MERMELSTEIN and EDWINA MEYER, Hofstra University

The objectives of this study were (1) to ascertain the effectiveness of various conservation training techniques, and (2) to ascertain whether some training techniques are more effective with certain populations than other training procedures.

Four training techniques, Cognitive Conflict, Language Activation, Verbal Rule Instruction, and Multiple Classification, all designed to induce the concept of conservation of substance, were utilized over three samples each from a different population. It was hypothesized that the cognitive conflict technique would induce the concept of conservation of substance over all three samples, whereas the other training techniques would not induce the concept of conservation regardless of the sample utilized.

In order to test the hypotheses, three samples of children between 3 and 6 years of age from the Long Island, New York, area were selected for the study. In particular 96 children (5 classes) age 3½-5 years, from the Hofstra University Nursery School, 100 Project Headstart children (5 classes) age 4-5 years, and 220 kindergarten children (8 classes) age 5-6, were selected for the training programs.

Within each of the 18 classes approximately equal number of males and females were randomly assigned to the four training conditions and one control condition. Half of the classes within each school district were then randomly selected for the pretest. All the subjects were post-tested three weeks, two-and-a-half months and five months after training.

Seven graduate assistants (two males and five females) were trained and were randomly assigned to the schools and classes.

In order to test the hypotheses, Dunnet's multiple comparison procedure was used. To examine the effects of pretesting, time, training, and the interactions relevant to the four hypotheses, a 2×5×3 analysis of variance design for partly correlated data was employed. A 5×3 analysis of variance design was employed to examine the effects of time and training for the specific populations.

The results indicate that the congitive conflict hypothesis is rejected whereas the other hypotheses are supported. With the exception of the overall effect of time in both analyses of variance, all other factors and their interactions were not significant at the .05 level.

It is concluded that the "Piagetian" concept of conservation of substance, as measured by the specific criteria described, was *not* induced by a variety of training procedures with various populations. Further it is suggested that language interferes with rather than facilitates acquisition of the concept of conservation of substance.

◆ ◆ ◆

Learning and the Conservation of Substance

J. S. CARLSON, University of California

The effects direct tuition may have on Piagetian concept attainment is unclear. Though evidence does exist which supports the notion that training can be efficacious in moving the child from "non-conservation" to "conservation," the mechanisms that may be involved have not been isolated.

In order to investigate further the effects of direct tuition on the conservation of substance, the following was done.

Sixty-five children (mean age, 6; 7) were randomly assigned to four experimental groups and one control group. The independent variables were (1) levels of verbal instruction (highly and minimally verbal), and (2) activity (direct experience and demonstration). The dependent variable was a conservation of substance test administered immediately after training and one month later. All training and testing was individual.

The results may be summarized as follows:

1. All experimental groups outperformed the control group on the first post test.

2. All experimental groups outperformed the control on the delayed post test with the exception of the group receiving minimally verbal x demonstration training.

3. On both the first and delayed post tests, all training methods used were equally efficacious.

4. Analysis of the difference scores between the first and delayed post tests showed no changes.

The results are discussed with reference to the development of mediational processes, and more than a "pseudoconcept" of conservation of substance was formed.

◆ ◆ ◆

Training Effects and Concept Formation: A Study of the Conservation of Continuous Quantity by Children

JOHN O. TOWLER, Purdue University and L. D. NELSON, University of Alberta

The purpose of this study was to investigate the development of children's conceptions regarding the conservation of continuous quantity, and the effects of training on the acquisition of this concept. Piaget's classical pretest of conservation of continuous quantity was administered to a stratified random sample of one hundred grade one subjects who were subsequently classified according to their knowledge of the conservation principle.

Subjects who were deficient in their understanding of the concept were randomly assigned to experimental and control groups. Experimental groups of ten subjects each participated in a short training experience while the control groups received no training. After the training session (or an equal interval of time in the case of the control group) each subject was retested with the classical pretest. In addition, the experimental subjects were given a posttest after a three week interval. This test was designed to assess the subjects' ability to retain the conservation concept, and to ascertain whether they could transfer the effects of their training to a new situation.

The major hypothesis of this study was that learning would take place if certain crucial aspects of the conservation principle were presented to the child in a manner which he could understand. These crucial variables were hypothesized to be: the understanding that a fluid retains its identity despite transformations in shape, an understanding of the compensatory relationships between height and width, and lastly, an insight into the principle of reversibility.

The major findings were as follows:

1. The training procedures were highly successful and resulted in helping a significant number of the subjects acquire an understanding of the conservation principle.

2. With the exception of one subject, all the members of the experimental group were able to retain the learning they had acquired in the training session, and were able to apply this knowledge to a new situation three weeks after the training despite an extinction suggestion made by the examiner.

3. No significant changes took place in the control groups' understanding of the conservation principle during the time in which the experiment took place.

2:00 to 2:55 P.M. **SESSION 6.2** **DIVISION C**

TEACHING THE DISADVANTAGED I

CHAIRMAN:

Frank Garfunkel, Boston University

PAPERS:

Comparison of Achievements in Special Reading Classes Using Guidance, Skill-Content, and Combination Approaches

JOSEPH M. LILLICH, Indiana University

The purpose was to determine whether significant differences existed between scores obtained by pupils with reading difficulty on the study's evaluative instruments.

Testing of hypotheses involved analysis of variance and subsequent t tests.

The guidance approach identified factors of emotional and social adjustment, and provided special activities to aid pupil growth. Reading skills typically included in special readng classes were not recommended for inclusion, nor were they identified for direct teaching.

The skill-content approach involved activities aimed at teaching directly the traditional skills used in reading.

The combination approach combined the guidance approach and the skill-content approach with time shared equally.

Samples were located in six elementary schools. The total sample consisted of 360 children from grades four, five, and six. The *California Reading Test* and the *Gray Oral Reading Test* were used to determine reading levels and achievement. These tests were administered on the first, sixtieth, and one-hundred-twentieth day of the study.

Null hypotheses were rejected by F Ratios significant at the .01 level of confidence. The subsequent t tests found that on the *California Reading Test* and *Gray Oral Reading Test* there was a significant difference at the appropriate evaluation intervals between the mean gains at the .01 level of confidence in favor of the guidance approach over the skill-content approach, and the combination approach over the skill-content approach.

Results on the *Gray Oral Reading Test* showed a significant difference at the appropriate evaluation intervals between the mean gains at the .01 level of confidence in favor of the guidance approach over the combination approach.

Focusing on guidance provides experiences for pupils which lead to initially higher performance on reading tests within this study. Initial "freedom to try", fostered by the guidance approach, could be profitably followed by direct teaching of reading.

Rehabilitation of School Dropouts: Results of a Follow-up Study

JOHN C. EGERMEIER, Oklahoma State University; VICTOR O. HORNBOSTEL, Bowling Green State University; J. PASCHAL TWYMAN, Tulsa University; GAYLEN R. WALLACE, U. S. Office of Education

This is a report of a 44-month investigation in which three curricular approaches to rehabilitating school dropouts were evaluated under experimental conditions. The project represented an attempt to bring together the resources of many agencies to seek solutions for a persistent social problem. Involved in the project at various stages were a major private foundation; a metropolitan public school system; a state university; and local, state, and federal governmental agencies.

Individuals who participated in the experimental project received a combined program of academic and vocational training; only vocational training; or only academic training. A control group made up of comparable individuals who received no training was also included. Over 300 school dropouts between 17 and 22 years of age participated. Approximately 50%were members of minority groups, including individuals of American Indian, Negro, and Spanish-American extraction.

The design of the study permitted evaluation of achievement during training and also of vocational success, continuation of education, and general satisfaction with life during a two year follow-up period.

In no case during the follow-up period was a significant difference found between groups that had received an experimental treatment. In comparing data from the three groups that had received training with data from control group subjects who had received no training, a number of statistically significant differences were found. Fewer differences were found in the second year than in the first year following training.

On the basis of the measures used to assess job satisfaction, employer ratings, and general satisfaction with life, no significant differences were found among the subject groups in this study. Similarly, there was no significant difference in the proportions of subjects in the various groups who continued their education or training during the two-year follow-up period. Specific findings, limitations, conclusions, and recommendations will be reported.

◆ ◆ ◆

The Effectiveness of Emphasizing Reading Skills In an English Course for Underachievers

ROBERT EMANS, Temple University, and RAYMOND URBAS, Wisconsin State University—Stevens Point

The purpose of this project was to study an adjusted English curriculum designed for children who were not achieving in accordance with their intellectual potential. It investigated the effectiveness of emphasizing reading skills as part of an English curriculum with a consequent de-emphasis in English content. The project had the following hypotheses:

1. Emphasis on reading for underachievers will be related to greater achievement in reading skills.

2. The consequent reduction of grammar and essay writing as a result of an emphasis on reading skills will not significantly impede achievement in grammar and essay writing as compared with similar groups of children.

3. Children in a course of study adjusted to their needs will express more positive attitudes toward learning, than children who are in a course of study that is not so adjusted.

The following procedures were used:

1. Three matched groups of children who were not achieving in accordance with their intellectual potential were formed. One group was given instruction in reading skills. Another group was given English instruction adjusted to the academic underachiever, including formal grammar and essay writing. The third group was included within regular English classes.

2. Pre- and post-assessments were made in reading, grammar, essay writing, and attitudes.

3. Three groups were compared in reading, grammar, essay writing, and attitudes after one academic year.

Hypothesis I and Hypothesis II were accepted in light of the results of this study. Not only was Hypothesis III rejected, but the opposite was found to be true. Therefore, the conclusion of this investigation is that an English course of study for the underachiever may be so adjusted as to help in the development of basic reading skills. The consequent de-emphasis in grammar and essay writing will not hurt the child's abilities in these areas. However, poorer attitudes toward school and learning may accrue as a result of such an adjusted curriculum.

◆ ◆ ◆

Preschool Intervention Through a Home Teaching Project

DAVID P. WEIKART and DOLORES Z. LAMBIE, Ypsilanti Public Schools

A preschool intervention project was conducted in the homes of 35 culturally disadvantaged families, to determine the feasibility and immediate effectiveness of home teaching as a method of improving the cognitive growth of four year old disadvantaged children. Matched experimental and control groups were established using cultural deprivation level of the family and Stanford-Binet IQ of the child as criteria. Various environmental, demographic, and attitudinal variables were collected through interviews and observations. The basic educational procedure of the project was a one-and-a-half hour per week home visit to each participating family, during which a carefully individualized educational program involving the mother and the child was offered.

The experimental group obtained a statistically significant greater change score on the Stanford-Binet than did the control group. On the whole, the home teaching program had the effect of 'freeing' the experimental group children from the environmental determiners of intellectual growth. Weekly ratings of the mother were made by the teacher working in the home. High IQ gain ($\overline{X} = 18.9$) was associated with qualities in mothers that teachers regarded as 'bad' (such as shallow, insensitive). Also, high IQ gain children tended to come from families who were 'worse off' in terms of general environmental variables. The environment created in the home by the mother is a critical determiner of intellectual functioning.

2:00 to 2:55 P.M. **SESSION 6.4** **DIVISION C**

PROGRAMED INSTRUCTION I

CHAIRMAN:

William A. Deterline, General Programmed Teaching

PAPERS:

Student Use of Behaviorally-Stated Objectives to Augment Conventional and Programmed Revisions of Televised College Economics Lectures

PHILIP W. TIEMANN, University of Illinois at Chicago Circle

Two steps of the programmed instruction process were investigated using criteria of student achievement and attitude. Objectives implicit in eight videotaped lectures, which had been revised using intuitive, conventional procedures, were identified as "general" objectives. As a first step, these general objectives were defined in terms of observable performance standards expected of students. The second step, validation, consisted of tryout and revision of each lecture until representative students attained the standards imposed by the "specific" objectives. Two groups viewed programmed revisions, and another two groups viewed conventional revisions. In each case, one group was provided with general and the other with specific objectives. Initial proficiency was pretested; learning was analyzed with criterion-referenced posttests covarying on pretest scores. A standardized questionnaire assessed six attitude factors.

Greater preferences for more objectives, greater indicated use of objectives and more favorable attitude were associated with the provision of specific objectives. ANOCOV of a 50-item multiple-choice (midterm) immediate posttest disclosed high scores to be associated with viewing programmed lectures ($p < .05$) but not with specific objectives treatment. Similar ANOCOV of 25 items repeated as part of the final examination disclosed high scores to be associated with the specific objectives treatment ($p < .05$) but not with programmed lectures. Other data suggested the assumption that the close association between objectives and criterion-referenced tests was not perceived until the mid-term posttest which structured the contingencies for subsequent use of objectives.

Precise contingencies appear to be needed when specific objectives are provided to guide independent study. If specific objectives can be provided and contingencies structured, the data suggest the investment required to validate instruction against objectives might have to be justified on some basis other than enhanced learning and attitude.

The Effects of Prompting in Programmed Instruction as a Function of Motivation and Instruction

GERALD W. FAUST, University of Illinois

An experiment was completed to investigate the conditions under which permitting the possibility of correct responding on the basis of inadequate inspection behavior is detrimental to performance in a programmed instruction lesson. Each of two versions of a stylized program to teach Russian vocabulary was given to college students. The frames of one version contained underline prompts which were hypothesized to permit *S*s to respond correctly without attending to the entire frame. The second version contained no such prompts. Prior to receiving the training programs *S*s received treatments intended to modify their motivation toward the learning task. In addition they received instructions which were intended differentially to affect the amount of time spent on the task. As had been hypothesized, under conditions of low motivation or time pressure the value of the attentional control inherent in the no-underline programs was accentuated.

◆ ◆ ◆

The Interaction Between Response Mode and Response Difficulty In Programmed Instruction

CHARLES KARIS, JAMES E. GILBERT, ALVIN KENT, Northeastern University

The inconsistency among experimental findings regarding the effectiveness of overt vs. covert responding in programmed instruction has been attributed by a number of reviewers to differences among programs in response difficulty, with the more difficult response favoring the overt mode. However, because different programs have been used to represent varying levels of difficulty, questions concerning the influence of such uncontrolled variables as style, presentation techniques, subject's interest in content, etc., have confounded this conclusion.

To test the hypothesis concerning the interaction between response mode and difficulty, a 196-frame linear program on medical terminology was developed which allowed for the determination of the effects of two levels of response difficulty within the same program. A statistical analysis of program errors and post-program retention scores demonstrated that responses requiring the reproduction of medical terms, as opposed to responses requiring the definition of medical terms, represented the more difficult responses. Accordingly, 50 college students were randomly assigned to an overt response group and a covert response group, and the program was administered over a two-day period. One week after the program each student was given a retention test measuring proficiency in the recall of (a) medical terms, and (b) definitions of medical terms.

For analysis, retention test responses were delineated into categories ranging from perfectly correct responses to varying degrees of approximation to the correct responses. The only statistically reliable difference obtained was on the recall of medical terms, and this only when *perfectly correct* responses were compared. As the degree of approximation to the correct response decreased, the differences between the two groups became smaller.

These findings indicate that the retention of programmed material is more effective in a program requiring an overtly constructed response only when the retention test necessitates the perfect reproduction of relatively difficult material.

◆ ◆ ◆

The Effect of Familiarization and Response Mode on the Programmed Learning of Foreign Language Vocabulary

WILLIAM B. GILLOOLY, Rutgers, The State University; and JAMES FURUKAWA, The Johns Hopkins University

A constructed response (CR) mode in programmed instruction has been shown to be more effective than a multiple choice (MC) mode when (1) performance is measured with a CR, not an MC, criterion test; and (2) when novel (i.e., unfamiliar) terms are involved. This study investigates whether familiarization-like procedures intrinsic to the CR mode account for CR mode superiority.

Programmed materials designed to teach the meanings of 16 Hawaiian words were constructed so that each of five information frames was followed by question frames in either a CR or MC format. Extrinsic familiarization materials preceded the programmed lesson, and consisted of two sheets on each of which were typed three times either the 16 Hawaiian words encountered later in the program (the Relevant Familiarization—RF) or 16 other Hawaiian words (Irrelevant Familiarization—IF). *S*s pronounced each word three times while writing it. The four possible packets of materials (RF/CR; RF/MC; IF/CR; IF/MC) were assigned randomly to 132 college Ss (33/condition) yielding a 2×2 factorial design. Following the programmed lesson, Ss were given three tests (CR, MC, and spelling). The hypothesis that there would be a familiarization by response mode interaction (that is, that CR-MC differences, especially on the CR test, are altered by familiarization) was tested by means of ANOVA.

Although there were significant main effects due to familiarization and response training mode on both the CR (RF>IF; CR>MC) and the MC test (RF>IF; MC>CR), the hypothesized interaction failed to materialize. On the spelling test, only the main effect due to familiarization was significant (RF>IF).

Instead of supporting the hypothesis that familiarization is responsible for CR mode superiority on a CR test, these results support a response-continuity hypothesis (that the CR mode is superior training for a CR test and the MC mode is superior training for an MC test regardless of familiarization). The results also attest to familiarization's efficacy as a pedagogical technique, for the RF group was superior on all three tests regardless of the response training mode employed.

2:00 to 2:55 P.M. **SESSION 6.5** **DIVISION D**

JOINT SESSION WITH NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION

RESEARCH ON SPECIAL GROUPS

CHAIRMAN:

Calvin O. Dyer, University of Michigan

PAPERS:

Reliability Coefficients of Certain Tests Used in the San Antonio Language Research Project

RICHARD D. ARNOLD, The University of Texas at Austin

Research dealing with cultural deprivation has precipitated concern over the reliability of measuring instruments used with various subgroups within the disadvantaged population. In particular tests standardized on middle class Anglo-American samples have been questioned. This report is an outgrowth of studies conducted to assess the effects, in terms of reliability, of using tests with a population other than that for which they were designed.

The sample upon which this investigation is based consists of over 200 Mexican-American children selected randomly from third grade classes of several San Antonio, Texas, elementary schools. It consists of children from a densely populated area where the median family income is less than $3,000. Many of the children speak little or no English when they enter the first grade.

The general questions are:

(1) If a test standardized on middle class Anglo-American pupils is applied to this disadvantaged population, what will be the effect upon the test's reliability?

(2) In terms of reliability, how does such a test compare to similar instruments developed for this type of population?

The tests investigated were the Metropolitan Achievement Tests, the Inter-American Reading Tests which were designed for Spanish-speaking children, and the IPAT Culture-Fair Intelligence Test. Cronbach's alpha was employed as the estimate of reliability.

Reliabilities computed for the reading subtests and total score of the Metropolitan were quite high (.88-.95) and compared favorably with the published reliabilities. The coefficients for the Inter-American Tests administered in English as well as in Spanish were also fairly high (.51-.91). The IPAT Culture-Fair Intelligence Test's reliabilities (.21-.77) were somewhat lower than expected, perhaps indicating that the test is less "culture fair" than its name would imply. All tests, however, were found to possess sufficient reliability so that at least the total scores may be used for group comparisons.

Racial Differences in Inheritance Estimates for Tests of Spatial Ability

R. T. OSBORNE, University of Georgia, and A. JAMES GREGOR, University of California, Berkeley

In order to examine racial differences in the degree of genetic control of perceptual motor tasks a sample of 284 pairs of like-sexed twins was administered a battery of eight tests of spatial ability.

The sample consisted of 172 pairs of identical (MZ) twins and 112 pairs of fraternal (DZ) twins. There were 121 boys and 163 girls; 43 pairs were Negro and 241 were white. The age range was from 13 to 18.

The zygosity of the twins was established by serological tests performed under the supervision of Mrs. Jane Swanson of the Minneapolis Ware Memorial Blood Bank. The following factors were tested for "A, B, O, M, N, S, s, P_1, P_2, Rho, rh′, rh", Miltenberger, Vermeyst, Lewis, Lutheran, Duffy, Kidd, Sutter, Martin, Kell, Cellano, and occasionally some others. Twins that differed on one or more of these serological types were considered dizygous.

One hundred and twelve pairs of twins were diagnosed as definitely DZ, that is, they differed on at least one independently inherited blood group. Using only the results of serological tests the remaining 172 pairs were diagnosed MZ with the probability of accuracy no less than 95 times in a hundred.

Three different heritability ratios were computed:

1. Holzinger heritability coefficient was computed by the formula: $H = (r_{mz} - r_{dz}) / (1 - r_{dz})$

2. A heritability ratio proposed by Nichols (1965) was computed by the formula: $HR = 2\,(r_{mz} - r_{dz}) / r_{mz}$

3. The F ratio for testing the significance of the difference between the within-set variance of MZ and DZ twins was computed by the formula: $F = \partial wdz^2 / \partial wmz^2$

All of the MZ correlations were greater than the corresponding r's for the DZ twins.

The remarkable agreement among three heritability ratios for the Negro and white subjects invites speculation that there are no racial differences in the degree of genetic control for the mental abilities represented by the eight spatial tests used in this investigation. It is further observed that the abilities represented are independently inherited with as much as 78 per cent of the within-family variance accounted for by hereditary factors. Heredity and environment produce significantly greater differences in fraternal twins on mental tests of Visualization, Perceptual Speed, and Spatial Orientation than environmental influences alone produce in identical twins.

Differences between Anglo and Non-Anglo Children on Factorial Dimensions of School Anxiety and Coping Style

BEEMAN N. PHILLIPS, The University of Texas at Austin

Socio-cultural status differences in anxiety have been reported, although this difference may be influenced by differences in defensiveness in such groups. Also, sex differences in both anxiety and defensiveness are frequently obtained. In a recent USOE supported study of school anxiety, coping style, and various aspects of school behavior, differences in school anxiety and coping style were obtained for Anglo (white) and Non-Anglo (Negro and Mexican-American) groups of elementary school children. School anxiety and coping style, which were initially developed through factor analytic techniques (Kaiser, image analysis), were further factor analyzed; and four replicated school anxiety, and six replicated coping style factors were obtained using data obtained in both fourth and fifth grades. These factors were, for school anxiety: fear of negative valuation by others, fear of taking tests, lack of confidence in meeting expectations of teachers and parents, and physiological reactivity associated with low tolerance of stress. For coping style these factors were: seeks good relations with peers, seeks recognition of peers, willingness to admit negatively perceived affect, willingness to admit negatively perceived motives, seeks recognition of authority, seeks good relations with authority. Fixed-model analyses of variance, with sex and socio-cultural status as the fixed effects, were performed; and generally significant sex and socio-cultural differences were obtained, including a number of significant interactions, with Non-Anglos and girls generally being higher on the dimensions of school anxiety. Implications of specific results, e.g., for fear of taking tests, are discussed both in terms of theory and educational relevance. Results for dimensions of coping style are integrated into a "self-enhancing" tendency rationale.

◆ ◆ ◆

Word Associations to Selected Form Classes for Children Varying in Age and Intelligence

CHARLOTTE WILLIAMS and M. H. TILLMAN, University of Georgia

Recent studies of word associations have suggested that the developmental trend from heterogeneous to homogeneous responding is highly related to the acquisition of syntax. The present investigation was designed to study the effects and interaction effects of selected subject characteristics on homogeneous word associations to certain form classes. The specific research objectives were: (1) to examine the role of intelligence level in children's word associations for selected form classes; (2) to examine the role of age level in children's word associations for selected form classes.

Subjects were retarded, normal, and superior children between ages 4-0 and 11-11 grouped by two-year age spans. Two tasks were used to examine children's utilization of the English parts-of-speech: a word association task (Brown and Berko, 1960) consisting of 36 high-frequency stimulus words representing six selected parts-of-speech categories; a word usage task consisting of nonsense syllables introduced within a syntactical context in such a manner that the six form classes of interest were represented.

Results were as follows: (1) trends over CA levels were present—significant linear trends occurred in every instance; quadratic and cubic trends varied with the task, form class, and intelligence level of the subject; (2) retarded, normal, and superior groups showed differences and similarities in number of homogeneous responses—relationships varied with the task, form class, and age level of the subject.

These results generally support the following conclusions: (1) homogeneous responding increases in a linear fashion across age with the easier form classes leveling off in some instances: (2) homogeneous responding to form classes follows a regular sequence of development regardless of intelligence group; (3) performance level differs when retardates are compared with either or sometimes both the normal and superior groups while performance level does not differ when normal and superior groups are compared; (4) rate of homogeneous responding appears to be similar for all intelligence groups on the more common form classes while the groups respond differentially in some instances on the more complex form classes.

2:00 to 2:55 P.M. **SESSION 6.6** **DIVISION E**

BEHAVIORAL COUNSELING

CHAIRMAN:

Ernest Spaights, University of Wisconsin, Milwaukee

PAPERS:

The Effect of Group Counseling on the Behavior of Educable Mentally Retarded Boys

PHILIP H. MANN, University of Miami (Milton D. Jacobson, University of Virginia)

Emphasis placed on intellectual factors by many educators has resulted in only parenthetical interest in the personal and social adjustment of the retarded child, though this adjustment is important, and may be beneficial to learning.

In most schools, counseling services are limited to normal children, and are conspicuously lacking for special children. This is a questionable practice.

This study attempted, with group counseling procedures, to change self concepts of educable mentally handicapped boys with behavior problems. It also investigated the effects of counseling on the academic achievement, anxiety, deportment, and intelligence of these boys.

36 boys with ages ranging from 9-13 and I.Q.'s ranging from 56-80 were divided into control and experimental groups which were matched on chronological age, intelligence, and socio-economic status.

The experimental group received 12 one-hour counseling sessions from the principal investigator in a structured but supportive environment.

The control group spent equal time with the same counselor in a noncounseling, but structured, library-study situation.

Pre- and post-measurements were given and were subjected to analysis of variance and covariance.

The experimental group exceeded the control group in self concept improvement, anxiety reduction, academic achievement, and deportment (.05 level of significance).

This study, though essentially exploratory, demonstrates the benefits which educable mentally handicapped boys can obtain from group counseling and its effect on learning.

A Comparison of the Reinforcing Value of Five Stimuli Conditions on the Production of Self-Reference Statements in a Quasi-Counseling Situation

JOHN J. KENNEDY, University of Tennessee

This study was designed to determine the reinforcing value of five commonly employed stimuli in a setting resembling that of an initial counseling interview. The five stimuli conditions studied were: (1) "mm-hmm"; (2) "mm-hmm" accompanied by an affirming head nod and smile; (3) "good"; (4) "I see"; and (5) a short paraphrase. The stimuli were verbalized by Es, and self-reference statements emitted by Ss constituted the verbal response class to which the stimuli conditions were applied and evaluated. There were, in addition, two secondary areas of descriptive exploration. The first concerned the relationship between Ss' conditioning performance and reported awareness of the response-reinforcement contingency, while the second involved a consideration of the relationship between Ss' free operant performance and subsequent conditionability.

The Ss consisted of 72 randomly selected females majoring in elementary education. Three counseling graduate students, trained within the context of a pilot investigation, served as Es.

Ss were engaged in an experimental interview situation consisting of four discrete segments: (1) free operant, (2) conditioning, (3) extinction, and (4) the administration of an oral test of awareness. Ss were instructed to "Talk about anything that they would like to talk about". After establishing a base rate of self-reference responding, the Es attempted to reinforce all self-references emitted by Ss by responding to the operants with one of the selected stimuli. Extinction of the operant behavior was subsequently attempted by withholding the selected stimuli condition.

Analyses of variance and X^2 procedures were used to analyze the data.

Analyses revealed that emission of self-reference statements can be increased by E responses in a conversational setting, and that certain stimuli are more effective than others. Specifically, the paraphrase and the neutrally toned "mm-hmm" proved to be maximally effective. The use of "mm-hmm" with affirming head nod and smile, and the "I see" response, were clearly ineffectual as reinforcers of the operant.

Further, it was found that successful operant conditioning of self-reference statements is possible in the absence of reported awareness, and that a substantial linear, or curvelinear, relationship between free operant performance and condition-ability, did not materialize.

Implications suggested by these findings for the evaluation of work previously undertaken in verbal operant conditioning and for the execution of future studies *will* be discussed.

Devising Social-Model Counseling Procedures for Elementary School Children

RAY E. HOSFORD, The University of Wisconsin

The general purpose of this study was to conduct an experimental test of two social model counseling procedures which can be used by school counselors to assist shy elementary students in learning how to participate in class discussions more effectively.

Recent research emanating from social learning theory has shown that observing models has profound effects on the behavior of observers. Studies have demonstrated that audio, video, and physically present social models are effective in influencing behavior.

If modeling procedures are to be utilized effectively by elementary school counselors, it is necessary that we determine the most effective way of presenting the social model. Specifically, this study sought to determine which of two types of model presentations (film or audio-tape) would best promote subject recall of those behaviors and suggestions which the social model sought to promote.

The type of model presentation did significantly affect the amount of content recall for girls, but not for boys. Girls exposed to the model counseling session presented on film, recalled a significantly greater number of model behaviors and suggestions than did girls exposed to the same social model presented on audio tape.

The type of model presentation affected the degree to which girls indicated that they would utilize the model's suggestions if they had such a problem. Significantly more girls after exposure to the film model presentation indicated that they would use the model's ideas, than did girls exposed to the audio-tape model treatment. Trends in the data indicate that the boys were similarly predisposed, but the difference did not reach traditional levels of significance.

Subjects exposed to the film model indicated significantly greater confidence in the counseling process than did subjects exposed to the audio-tape model.

Exposing subjects to two social model counseling sessions did significantly affect improvement of content recall if the second model presentation was on film. Subjects viewing the film model one week after hearing the audio tape demonstrated a significant improvement in the amount of content recall, whereas subjects hearing a second audio-tape model session one week later did not improve significantly the degree to which they could recall the content of the social model interview.

Physiological Changes During Muscle Relaxation Training and the Visualization of Fearful and Neutral Situations

JOHN M. GROSSBERG and HELEN K. WILSON, San Diego State College and University of Minnesota (T. Antoinette Ryan, Oregon State University)

Behavior therapies assume that maladaptive behaviors have been learned in the same manner as other behaviors, and can be eliminated in similar fashion. One widely used method, systematic desensitization with relaxation, has been applied to many handicapping fear reactions. Assuming that relaxation is antagonistic to fear, clients are trained to relax while progressively stronger fear stimuli are presented, until fear has been totally inhibited by the damping action of relaxation. The client is told to imagine fear situations, since it is too impractical to arrange progressive gradations in real life.

Our studies tested theoretical assumptions of reciprocal inhibition theory. Experiment I assessed the physiological effectiveness of relaxation training as compared with control procedures. If relaxation training does not reduce tension physiologically, this would undermine the assertion that relaxation inhibited fear arousal. Experiment II tested the assumption that visualizing fearful situations produces more tension than visualizing neutral situations. Unless this is so, it would be difficult to maintain that fear was extinguished by repeated visualizations.

Experiment I. Heart rate (HR), skin conductance (SC), and forehead EMGs of 10 male undergraduates who received muscle relaxation exercises, were compared with data from controls who received taped music, and another group who received no experimental stimulation. While groups did not differ significantly in HR, relaxation subjects showed consistent although non-significant decreases in SC and EMG.

Experiment II. 36 girls were first read and then told to imagine a fearful and a neutral scene four times while HR, SC, and EMG were recorded. For HR and SC, a) during reading there were no significant differences between fearful and neutral scenes; b) during imagining, fear scenes produced significantly more arousal; c) both measures declined significantly over trials. EMG failed to differentiate fearful from neutral scenes.

These studies generally support reciprocal inhibition theory assumptions that relaxation training lowers tension, and that visualization of fearful situations produces increased activation.

2:00 to 2:55 P.M. SESSION 6.7 DIVISION E

SELF-CONCEPT AND PUPIL PERFORMANCE

CHAIRMAN:

Elwyn H. Nagel, University of Iowa

PAPERS:

Opportunity as it is Related to Home Background and School Performance

FOREST HARRISON, Claremont Graduate School

The purpose was to investigate the relationship between student opportunity and home background, when school performance is considered. From each of six countries, four groups of thirteen-year-old students—advantaged-successful, advantaged-non-successful, disadvantaged-successful, and disadvantaged-nonsuccessful—were drawn from the pool of 29,311 students in the data bank of the International Project for the Evaluation of Educational Achievement (IEA). No countries other than the United States were identified, as was requested by the Standing Committee of the IEA.

Parents' educations and father's occupation were used to estimate home background, and performance was defined by mathematics achievement and an interest in mathematics index. Student opportunity was a teacher's rating of the opportunity which each student had been afforded to learn the content of the mathematics achievement test. In the analysis for each country, home background was crossed with performance; an analysis of variance for a nonorthogonal design was used to test these effects.

The results were: student opportunity was related to performance in five countries, including the United States, The successful students in these countries had been afforded greater opportunity than had the nonsuccessful students. In the remaining country, all four groups of students had been afforded the same opportunity.

Moreover, when the effect of performance was removed, opportunity was also related to home background in four countries, including the United States. The advantaged students in these countries had been given a greater opportunity to learn the mathematics content than had the disadvantaged. Within the other two countries, advantaged and disadvantaged students had been afforded similar opportunities. These results will be discussed in terms of the issues underlying this study.

Relations Between Self-concept and Indices of Performance in Mathematics

DAVID L. KLEMMACK and JAMES J. RYAN, Minnesota State Department of Education

Brookover and others have demonstrated that a pupil's self-concept of his ability in school subjects is positively related to classroom performance as measured by grade point average (GPA) even when differences in ability are controlled. This study was carried out to determine (1) the generality of self-concept-performance relationship for (a) a sample of pupils in ninth grade algebra and (b) more situation-free measures of performance, achievement tests, as well as GPA; and (2) whether sex and teacher are relevant variables affecting the nature of this relationship.

Premeasures of self-concept of ability in mathematics (SCAM) and of mathematics proficiency and post measures of performance on two somewhat different mathematics achievement tests and algebra grades, were obtained from a sample of 270 pupils in 14 algebra classes instructed by different teachers in different schools.

Regression analyses indicated that even with differences in initial proficiency, controlled SCAM was significantly positively related to performance on each of the two achievement tests as well as to grades, with a somewhat higher proportion of grade variance than test variance being accounted for by SCAM. An analysis of covariance with SCAM and preproficiency as covariates revealed significant sex and teacher differences on the performance measures, but neither teacher nor sex affected the SCAM-performance relationship.

The results indicate that SCAM is related to indices of achievement which are not as likely as grades to reflect attitudinal aspects of classroom performance, i.e., to indices which more clearly reflect knowledge acquisition, per se.

Similarities and differences between these and the results of other studies in this area will be discussed as will their implications concerning classroom evaluation and pupil factors involved in the relations between self-perception of academic ability and performance.

◆ ◆ ◆

Correlates of Birth Order: Effect or Artifact

ROBERT J. PANOS, American Council on Education

Recently researchers have been giving more attention to the effect of ordinal family position on the academic achievement and later vocational development of the student. Although a large number of studies indicate the pervasiveness of "birth order effects," the suggested causal relationships are often not testable in the context of the particular study, or are contradicted by other research. This result follows from the fact that birth order data are usually collected as ancillary information, without deliberate design, and are then correlated with other data. Perhaps the principal limitation of current hypotheses concerning the effects of birth order is that the data are not obtained in a context that permits testing alternative hypotheses. The principal objective of this study was to determine the correlates of birth order effects within a large national sample of entering college freshmen.

During the fall of 1966, birth order data were collected from over 250,000 entering freshmen students as part of the American Council on Education's Cooperative Institutional Research Program. A representative sample of 9,330 subjects was selected from the data file for this study. In addition to ordinal family position, the data included information with regard to age and sex distributions of siblings within family sizes, and information concerning the student's background, achievements, activities, and educational aspirations.

Preliminary analyses of the data indicated that, at a minimum, it would be necessary to introduce adequate controls for sex and for family size. Once these controls were included in the data analyses, observed birth order effects were negligible with regard to such variables as student achievements, school-related activities, and educational aspirations. The results of this study suggest that many of the hypotheses concerning "birth order effects" derived from the framework of psychological, physiological, sociological, and economic theories, may need to be reexamined.

◆ ◆ ◆

The Multiple Measurement of Self Concept

LEE S. SHULMAN, Michigan State University

As part of a longitudinal study of the vocational development of mentally handicapped adolescents, the self concepts of subjects were multiple measured using four independent methods of gathering self concept data. The four instruments used were (1) a 150-item self report scale; (2) an interview during which the subjects responded to questions concerning their perceptions of themselves physically, socially, academically, and vocationally; (3) a series of projective drawings; and (4) a five day series of level of aspiration measures which were gathered while the subjects were participating in a simulated work experience. The level of aspiration measure yielded two scores, the average level of aspiration ratio, which was the ratio of predicted production to previous production in the work situation and the Aspiration Index, which reflected the degree of optimism in the aspirations of subjects irrespective of the magnitude of the aspiration.

The study was based on the assumption that no single measure of self concept necessarily reflected subjects' actual perceptions of themselves.

Only by looking at the patterning of self concept relevant responses across modes of responding (cf. Campbell and Fiske) could an accurate estimate of the self conception of subjects be derived.

The data demonstrated that a small number of clear-cut patterns of responses were generated for the self concept variables. There was a clear difference between responses to self-report measures, i.e., the self-attitude scale and the interview, and performance measures, i.e., the level of aspiration ratio and the Aspiration Index. Differences in these self concept patterns were subsequently related to individual differences in the characteristics of the subjects. Finally, a transposed (Q-type) factor analysis replicated the patterns which had been identified by inspection.

Attempts to measure self concept have in the past been seriously weakened by the dependency of investigators upon single measures of self concept. The importance of multiple measurement of such personality traits in order to "triangulate" on the trait through use of contrasting assessment techniques must be explored further.

3:00 to 3:55 P.M. **SESSION 7.2** **DIVISION C**

MICROTEACHING

CHAIRMAN:

James P. Shaver, Utah State University

PAPERS:

The Effectiveness of Self-Instruction in Teacher Education Using Modelling and Video-Tape Feedback

DAVID B. YOUNG, University of Maryland

The purpose of this study was to appraise the relative effectiveness of various modes of training intern-teachers in specific skills without direct supervisor-teacher conference.

The experimental design provided for the investigation of three modelling protocols and focusing techniques and two self-feedback procedures. The modelling protocols consisted of a "specific illustration" model (a teacher stands before a camera without a class and gives discrete examples of the teaching skill) and a "complete" model (examples of a skill in a lesson context) presented with either a contingent or noncontingent focus.

The self-feedback procedures consisted of subjects viewing their own performance with a contingent focus (the experimenter reviewed the video tape and recorded comments contingent to the behavior on the second sound track before the training session) or a non-contingent focus on the video tape.

The study was conducted in a "regular classroom" context during the Fall Quarter 1967, in forty public schools throughout the San Francisco Bay Area.

In general, the most effective modelling protocol is a combination of the "specific illustration" model and a "complete" model with a contingent focus. However, each modelling protocol appears to be more effective for training for a specific variable than for all variables in general.

The "specific illustration" model and "complete" model *combined* were *significantly* more effective for teaching the "variables": writing on the chalkboard, all visual highlighting (total), and student examples.

Neither of the above modelling protocols produced significant differences when employed separately.

Subjects viewing their own performance with a contingent focus used a significantly greater number of visual highlighting variables (total), visual aids, teacher-student examples, and total of all verbal variables.

The combination of viewing a "specific illustration" model, a "complete" model with a contingent focus, and viewing one's own performance with a contingent focus proved to be significantly more effective than other combinations for the variables: teacher and student examples, visual highlighting (total), and writing on the chalkboard.

The results of this study suggest that the following areas need further investigation: (1) the effectiveness of each modelling protocol when employed separately; (2) the relative effectiveness of supervision in-person and self-instructional techniques; (3) the most effective modelling protocol for a specific teaching skill; (4) the effectiveness of different teaching skills in producing learning; and (5) the degree of transferability of learning from the situation modeled to new and different situations.

◆ ◆ ◆

The Effectiveness of Microteaching in the Preparation of Elementary Intern Teachers

WARREN KALLENBACH, Far West Laboratory for Educational Research and Development

All teaching candidates ($N = 40$) in the 1967 San Jose State College Summer elementary intern teaching program were randomly divided into two groups. One group, the Microteaching Group, participated in a Summer microteaching program on campus with no off-campus contacts with students. The other group, the Student Teaching Group, participated in a limited Summer observation and student teaching program. Both groups otherwise had the same Summer program.

Pre- and post-Summer lesson excerpts (five minutes each) were videotape recorded for each of the candidates and these were independently judged, double-blind, by each member of two teams of trained evaluators. The evaluators used the Stanford Teacher Competence Appraisal Guide (STCAG) to judge the teaching skills from the videotaped recordings.

A field follow-up by two independent teams of trained evaluators made both a Fall and Spring assessment of each intern teacher. One team used STCAG and the other used Instrument for the Observation of Teaching Activities (IOTA). Each team member judged independently.

No significant differences in teaching skills or competence appeared between the two groups at the end of the summer programs or developed in the assessments during the school year. Judges' findings were found to be moderately but significantly correlated with some few exceptions. One team of judges found the Microteaching Group candidates significantly less able in teaching skills ($p < .05$) at the beginning of Summer but this difference did not appear at the end of Summer. This might appear to favor the Microteaching Group who, although they spent eighty per cent less time in teaching activities, reached comparable levels of teaching skills with the more able candidates in the ten-week Summer program, and these levels were maintained at an equivalent or higher level of competence throughout the school year.

The major contribution of microteaching as compared to Summer student teaching in an intern program is in the time saved in teaching activities by the microteaching program—over eighty per cent in this study.

◆ ◆ ◆

An Evaluation of the Effectiveness of the "Micro-Teaching" Technique in the Training of Elementary School Teachers

THOMAS B. GOODKIND, University of Connecticut

The three general objectives of this study were (1) to develop and field test an adaptation of the "micro-teaching" technique for the on-the-job training of student teachers; (2) to determine the value of the small, portable videotape recorder in the training process; and (3) to obtain some quantifiable data to evaluate the "micro-teaching" concept in teacher training programs.

Forty students were selected from the University program, half of which were involved in "micro-teaching" with videotaping, and half without the videotaping. Each taught two short lessons of 4-8 minutes and two long lessons of 10-20 minutes to the elementary children in her student teaching class. The experimental group had their performances recorded on videotape for playback and evaluation with their University supervisors and cooperating teachers. The control group students had similar evaluation sessions, but without the addition of the videotape recordings. All students had the chance to reteach the same lesson after the evaluation session, followed by a second session.

The Medley-Mitzel Observation Scale and Record was used as the principal evaluation instrument as a systematic observational technique to record each trainee's classroom performance and behavior. The OSCAR was administered by University supervisors four times during the student teaching period to all students in the study. Anecdotal information and informal observations as well as student comments provided additional forms of data.

Preliminary evaluation of data has indicated that those students in the experimental group displayed (1) a greater awareness of specific personal habits and mannerisms; (2) a greater awareness of specific teaching acts and techniques, particularly of the non-verbal type; (3) greater insight into the activity and interrelationships of children within the classroom; and (4) a greater awareness of the problems of pacing in their instructional program.

Effect of Video Tape Recording Feedback On Teaching Behavior of Student Teachers

JOHN A. VOTH, University of Missouri

Does feedback from viewing one's video tape recorded classroom instructional activities result in changed teaching behavior? If so, which teaching behaviors change? How can these behavior changes be measured? How much and what kind of feedback is required? The purpose of this study was to ascertain the relationship between the opportunity for student teachers to observe, analyze, and "evaluate" their attempts at student teaching (as recorded on video tape) and the extent of change in their stuent teaching behavior as measured by (1) verbal interaction between teacher and pupil; (2) types of cognitive response elicitations; and (3) group discussion roles.

The type of video recording feedback used in this study resulted in a significant increase in variableness of verbal interaction. Where at least 3 to 4 hours of structured feedback per hour of recording was obtained: there was a significant increase in "thoughtful answers" by pupils; in "indirect teacher influence;" and in two general categories of unclassifiable verbal interaction. No significant increase in cognitive response elicitations or group discussion roles was found.

Questionnaire responses from both student teachers and cooperating teachers indicated video recording feedback was very helpful in "improving" and "understanding" teaching; however, the data reflected this "help" or change in only a few of many variables.

3:00 to 3:55 P.M. SESSION 7.3 DIVISION D

JOINT SESSION WITH NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION

RESEARCH ON TESTING I

CHAIRMAN:

Roger T. Lennon, Harcourt, Brace and World, Inc.

PAPERS:

Objective and Essay Tests: Do They Measure Different Abilities?. . .

GLENN H. BRACHT and KENNETH D. HOPKINS, University of Colorado

This study investigated the extent to which essay and objective tests measure different abilities in evaluating academic achievement.

At the begining of an introductory course in educational psychology, the STEP Writing Test and a questionnaire which surveyed the subjects' preference for type of test and perceived success on essay and objective tests were administered to 279 students, mostly college sophomores, in five sections. Both an essay and objective test over the same content were administered at the end of each of the first two instructional units. Each test carried equal credit and time allowance.

After completion of the course, the four instructors participated in an extensive anonymous evaluation of the essay questions from 100 randomly selected students. Each essay question was evaluated independently by each of the instructors.

The correlation between essay and objective tests corrected for attenuation (over 1.0 in every case) suggested that the essay and objective tests were not measuring different cognitive abilities or achievement, although the objective test is a more reliable measure of academic achievement.

Length of essay and penmanship correlated significantly higher with essay scores than with objective test scores even when scholastic aptitude and g.p.a. were held constant. The partial correlation also showed that length was more highly related to the essay score than any other variable.

Cumulative grade point average correlated significantly higher with the objective tests than the essay tests.

Performance on essay and objective tests did not differ in the extent to which each was related to preference for type of test, scholastic aptitude, writing ability, grades in freshman English and general psychology, academic major, class, sex, and age.

The correlation between essay questions from the same test was significantly lower when the questions were read independently rather than sequentially, illustrating a carry-over effect.

Associative Assessment of Academic Achievement

WILLIAM N. JENNINGS, ROBERT RAMOS, and WILLIAM S. VERPLANCK, University of Tennessee

A new testing method requiring recall as do essays—and enabling computerized scoring methods, is presented as an alternative to the multiple choice examination. This test uses word associations for assessing academic achievement, that is, to determine how much has been learned.

Test format and administrative details of the testing procedure are described, and a report of the tests' psychometric parameters provided. The experimental examinations required college students to give four relevant word associations to a number of stimulus terms representing key concepts in the subject studied. Several formats and instructions have been evaluated, and a number of scoring methods investigated. Changes in both quality and quantity of associations with learning have been measured by giving pre- and post-learning tests to a single group, and also by presenting the same stimulus terms to introductory psychology students and to beginning graduate students in psychology.

Reliability has been estimated for several samples using both split-half and test-retest procedures; measures have been high, ranging from .724 (N = 25) to .943 (N = 28). The most stable estimate is .844 (N = 133). Tentative validity measures have used two different types of criteria. Association scores have been correlated with overall grade point average (G.P.A.), American College Test scores, position on the Dean's List, and also with multiple choice and short-answer examinations over the same subject matter. Correlations range from .356 with G.P.A. (N = 133) to .773 with short-answers (N = 25). Item difficulty and reliability of item difficulty have also been assessed.

The distributions of associations given by groups differing primarily in instructor have been compared. These comparisons indicate the exceptional sensitivity of these examinations for differences in what students learn from different instructors.

Some possible imitations of this form of examination are pointed out. How are scores on these tests related to scores on essays on the same content? Perhaps a relationship between these scores and 'verbal fluency' obscures the relationship between the number and accuracy of associations given, and factual grasp of the subject. Finally, we need to assess the relationship among the concepts covered in the word associations. Grasp of conceptual relationships may or may not be highly correlated with the simple number of associates given.

Relations between Achievement Test Item Characteristics as Judged by Teachers

JAMES J. RYAN, Minnesota State Department of Education

Teachers are closely involved in all aspects of the development, selection, use and interpretation of achievement tests. This often requires evaluations and judgments concerning specific test tasks (items), as well as making inferences from pupil performance thereon. There has been little systematic investigation of test item characteristics relating to or affecting these judgments.

This study was carried out to determine the nature of the relations among test item characteristics such as difficulty, discriminabiliy, relevance and overall quality as these are perceived by teachers when obtained with reference to instructional conditions and outcomes in two separately instructed classes.

For a sample of 60 secondary mathematics teachers, graphic scale judgments concerning relevance to the subject matter, difficulty and discriminability and overall quality for evaluation were obtained for each of 25 items in each of two math achievement tests with respect to the instructional content and performance of two of their mathematics classes.

One test was a standardized achievement test, the other a test developed to assess the instructional objectives common to several alternate programs at each grade level. Judgments were obtained on a standard form with a detailed characterization of the task.

For each teacher, within each class by test judgment situation second-order partial correlations were computed (across items) between each pair of scale judgments with judgments for the alternate scales partialed out. The proportions of teachers for whom each of the resulting six correlations was significant ($p < .05$) was determined, as was their relative magnitude within teachers. The judgment of overall item quality was significantly more frequently related to that of relevance than to judgments of difficulty or discriminability. The overall quality—difficulty relationship varied in frequency and direction between the two tests as did the generally high frequency of difficulty—discriminability relationships.

The results indicated that teachers tended to shift their judgment basis with more general properties of the test and that under some conditions there tended to be two relatively independent underlying dimensions of judgment.

The implications for the dimensions of teacher evaluation of pupil performance and for methodological aspects of test and item judgment are discussed.

The Influence of Test-Wiseness Upon Achievement. .

MERLIN WAHLSTROM, The Ontario Institute for Studies in Education and FREDERIC J. BOERSMA, University of Alberta

The purpose of the study was to show that a test-taking strategy can be taught by classroom instruction and that the knowledge gained will subsequently enable a test-wise examinee to obtain a higher score than an equally knowledgeable examinee who lacks test sophistication. An attempt was made to incorporate procedures for meeting the criticism directed at the Moore, *et al.* (1966) study while using more meaningful material to construct the criterion tests.

The results indicate that test-wiseness principles can be taught in a classroom setting and lead to improved achievement test performance. The results are intentionally magnified since the experimental groups were taught how to write items with common faults and then administered a test with poorly constructed items. It should be noted that all Ss had written many multiple-choice tests over the previous five year period which should have resulted in some gain in knowledge about test-wiseness principles, and that only the experimental group produced the anticipated gain.

While "good" items may be used to control for error variance associated with test-wiseness, the writers contend that teacher-made achievement tests contain items with faults, and that test-wise Ss often receive higher scores that Ss of equal intellectual ability. Moreover, it makes sense to assume that degree of knowledge about content material and ability to apply test-wiseness principles are highly related. Teachers should be made aware of the unwanted error component in test scores introduced by test-wise students, and students should be taught principles of testwiseness. Granted, the teaching of test-wise principles will not remove all variation in test scores due to differential test sophistication, but it will, in all probability, reduce this source of error variance in observed scores.

3:00 to 3:55 P.M. **SESSION 7.4** **DIVISION D**

DEVELOPMENT OF EDUCATIONAL MEASURING INSTRUMENTS

CHAIRMAN:

Warren G. Findley, University of Georgia

PAPERS:

The Construction of an Objective Measure of Self Concept.

M. P. GILL and V. R. D'OYLEY, The Ontario Institute for Studies in Education

The present study was designed to: (1) develop a theoretically oriented and objective instrument purporting to measure the self concept of high school students in an academic situation, and (2) to determine the factorial structure and predictive validity of such an instrument.

The final instrument consisted of 65 inventory-type statements. The instrument was administered to 1424 ninth-grade students from five collegiates. Each student was required to rate himself on a four-point scale on each of the items.

The stability coefficients for the self concept scale were .69 for boys ($N = 41$) and .60 for girls ($N = 27$). Hoyt's analysis of variance technique yielded reliability coefficients of .89 and .91 for boys ($N=227$) and girls ($N = 187$) respectively.

To provide a statistical basis for grouping of items that would permit psychologically meaningful interpretation of the data, students' responses to the self concept scale were subjected to principal components analysis. Eight interpretable factors were identified. The most important factors were achievement-related characteristics, acceptance by peers and teachers, and self confidence.

Validity coefficients for the instrument were computed using final average marks as the achievement criterion. The resulting coefficients were .42 for boys ($N = 674$) and .35 for girls ($N = 553$). Correlations with aptitude and IQ were .13 for boys and .09 for girls, for each of the predictors. The low magnitude of these correlations indicates that the new instrument has a great potential for tapping some dimensions of behavior not already covered by the most commonly used predictors of academic performance.

The Development of a New Test to Assess the Occurrence of Non-Standard English Among Disadvantaged Primary Children

JEROME TAFT, Dade County Public Schools

The Dade County Test of Language Development was designed to assess objectives for a Title I project under Public Law 89-10. The Special Reading Teachers Project, known in Dade County as PLATS (Project Language Arts Teachers), had as one of its objectives the increase of language development among educationally deprived children. To support the instructional program, it was considered necessary to identify occurrences of nonstandard language among linguistically handicapped students.

The test items reflect features which depart most radically from a standard English dialect. The following areas of language development were assessed: pronunciation, reception and grammatical factors of comparison, double negatives, double subjects, plurals, preterits, past participles, pronouns, uses of *do, be, have,* and subject-verb agreement.

Test I, Aural Reception, containing 30 items, was administered to 38 experimental classes (nineteen classes in grade 1 and nineteen classes in grade 2), during February and May, 1967. Four first-grade classes and five second-grade classes were used as control groups.

Ten classes were randomly selected from the original 38 classes tested. Among these ten classes, ten students were randomly selected from each class to take *Test II, Oral Usage,* the production phase of the test battery. Professional linguistic listeners were engaged to assess the tape recorded language production data. Production scores were quantified and recorded for each student. Tests I and II were administered on a pre and post basis.

Classroom teachers, working independently, assessed and rated each student who took Test II on specific production factors.

Means and standard deviations of the pupils by grade and sex were calculated from teachers' and linguists' ratings for both February and May performance. Pearson's product-moment correlations were computed.

Results indicate: A. After three and one-half months of instruction the experimental group scored higher on the post test than the control group. B. Generally, girls were rated higher ("higher" indicates better speech production patterns) than boys on the Teacher Rating Scale. Linguists also rated girls higher on the analysis of the tape recorded data; however, the sex difference was more marked at the second grade than at the first grade. C. Correlations between teachers and linguists revealed the following: 1. Correlations on February ratings between linguist and first grade teachers were high (.72) for girls and very low for boys (.13). 2. At the second grade the correlation for boys (.58) was higher than that for girls (.38) in a comparison of February ratings. 3. Correlations between ratings by teachers and linguists were consistently high in May for both sexes and grades ranging from (.53) for second-grade girls to (.79) for first-grade boys. 4. A comparison of pre and post teacher ratings for first and second-grade total groups correlated (.65) for grade 2 and (.16) for grade 1.

◆ ◆ ◆

Cognitive Preferences in Mathematics

KENNETH J. TRAVERS, University of Illinois; ROBERT W. HEATH, Stanford University; LEONARD S. CAHEN, Educational Testing Service

The development of an instrument, the Cognitive Preference Inventory, to measure student and teacher preferences for modes of expressing mathematical ideas was undertaken. Three a priori modes were specified: verbal, symbolic and graphic. Thirty items were written for pretesting. A different mathematics topic was used in each item. Directions for the student form of the test instruct the pupil to read the three response mode options in each item to determine the concept being dealt with, then choose the one mode which "appeals to him most." Directions for the teacher form of the test instruct the teacher to read the three options to determine the concept being dealt with, then choose that mode which he would prefer to use in teaching that concept to students.

The instrument was pretested using a sample of seventh-grade students ($N = 115$) and a sample of elementary school teachers ($N = 75$). Summary statistics indicate strikingly similar responses from students and teachers. In general, the symbolic option was most often preferred. The verbal and graphic scales form positively skewed distributions for both samples. The various reliability indexes, while apparently high, are not directly interpretable because of the ipsative nature of the three scales.

The Cognitive Preference Inventory appears to be equally reliable for students and teachers. The thirty-item test can be comfortably administered in a classroom period. A usable balance among preference scores is produced by the test. Cognitive preference, defined as "preference for mode of dealing with a concept" for the learner, and as "preference for mode of teaching a concept" for the teacher, appears to be characterized by similar patterns of preference in both students and teachers.

◆ ◆ ◆

The Development of Test Exercises for Assessing the Cognitive Skills of Elementary School Children

NORMAN E. WALLEN, ANTHONY H. McNAUGHTON, WILLIAM CRAWFORD and ENOCH SAWIN, San Francisco State College

One of the activities of the Taba Curriculum Development Project is the development of evaluation exercises or tests which may be used by teachers to assess the level of cognitive functioning of their pupils, particularly with Social Studies content. We have focused particularly on the following abilities: 1) ability to group discrete items in meaningful ways and attach an appropriate label to the groupings; 2) ability to form generalizations from data (with particular emphasis on decentering). Our approach has been to use brief stories—usually read to the children—followed by a series of questions and/or exercises. It seems necessary to establish a basic format of administration into which the teacher may insert her own content—in this case the particular story. The principal research problem then becomes one of developing a scoring procedure which is simple enough for ready use by teachers and which meets desirable technical standards.

Procedures have been developed for evaluating student responses along dimensions of accuracy, abstractness and tentativeness. They are appropriate for all elementary grades.

The materials have been tried with approximately ten teachers at each of grades one through six. Although some teachers were skeptical as to the childrens' ability to perform, the results allayed such concerns. Among the results of this project are: 1) Teachers commented that scoring the materials was of considerable value in improving their understanding of thinking skills. 2) The teachers' scores initially showed considerable discrepancy from scores assigned by the research staff but improved with practice. 3) Scores show expected relationships to age and socio-economic variables.

The exercises have also been administered, as part of our overall program evaluation, to approximately 700 sixth graders. In this instance administration and scoring was done by project staff. This project permits evaluation of the exercises 'at their best' in items of technical desiderata.

3:00 to 3:55 P.M. **SESSION 7.5** **DIVISION C**

SOCIOECONOMIC STATUS I

CHAIRMAN:

James T. Fleming, University of California, Los Angeles

PAPERS:

A Pilot Study of Children's Perception of Different Speech Forms

JAMES T. FLEMING, University of California at Los Angeles

The purpose of this exploratory study was to assess the degree of children's perception of socially significant speech variants in other children's speech, and to determine the extent of any stable relationship between children's and adult's perception of the same speakers.

Using a standard individual interview, two one-minute taped speech samples were obtained from lower-urban white (LUW), lower-urban negro (LUN), upper-urban white (UUW) and upper-urban negro (UUN) nine-year old boys. The eight speech samples were presented first to 25 adult listeners who, as expected, unanimously identified only the two LUN speakers; near unanimity of agreement was obtained for the two UUW speakers; LUW speakers were placed in all but the UUW category; UUN speakers were identified as upper-urban, with ethnic membership being split. The eight speech samples were then presented to 32 nine-year old boys selected approximately equally from a population of LUW, LUN, UUW and UUN boys. Listening individually, the S's task after each of the eight samples was to guess "who was talking" by pointing to one of four pictures which he had before him. The pictures depicted in stereotyped fashion (1) a white nine-year old boy against a typically slum background; (2) a negro boy of the same age against a background similar to (1); (3) a white nine-year old boy against a typically upper-urban (or suburban) background; (4) a negro nine-year old boy against a background similar to (3).

As with the adults, the two LUN speakers were the source of the most consistent S responses: in all but five instances, the 32 Ss identified both LUN speakers as negro, although some SES switching was apparent. The five exceptions were identical. LUN boys identified LUN speakers as LUW. Other trends were indicated, but more data and analysis are necessary. Despite the many acknowledged limitations of this study, preliminary findings encourage on-going refinement of instruments and techniques.

Cultural Differences in Reading and Understanding Geographic Maps

GAVRIEL SALOMON, The Hebrew University School of Education, Jerusalem, Israel (Joan E. Sieber, Stanford University)

The typical geographic map can be analyzed along different dimensions:

(1) The Semiotic Dimension: Maps use complex and varied symbolic codes, including replicas and symbols possessing varying relationships to the things for which they stand. Learning the code for one symbol may not aid in, or even interfere with, learning the code for another symbol.

(2) The Geometric Dimension: The map reader is expected to be able to understand certain spatial relations and the conservation of them (e.g., directions, horizontals, and verticals, etc.). Such expectations do not underlie non-Western graphical representations. Moreover, anthropological studies point out that such expectations are unjustified in children of non-Western origin. Learning difficulties could occur with such children.

(3) The Communicational Dimension: The map lends itself to hypothesizing and making inferences. From the symbols presented, additional information made on the basis of spatial relationships and combination of different items can be generated.

The reading of the map follows from the analysis of its characteristics and function. Three stages of reading it are hypothesized to exist:

(1) Identification of the signs and symbols. This requires certain learning strategies which can be, and are, specified.

(2) Analysis of spatial dimensions (distances, directions, and heights). These require the acquisition of certain spatial concepts.

(3) Interpretation of the mapped region; inferences from the map; grasping configurations of items. This requires an ability to differentiate the complex stimuli and integrate them into new patterns.

A comparison of two groups of children from two different cultural backgrounds was performed in Israel. The children ($N = 360$, seventh and eighth grades) were either from a Western or non-Western (so-called "culturally deprived") background.

A questionnaire accompanied by a map, one of Piaget's tests, and a test of drawing a map from a 3D model, were given to the subjects. Generally, the differences between the groups, with respect to the ability to read maps, increased from Stage I (identification) to Stage III (interpretation). The differences on all the tests were statistically significant favoring the Western-born children. Several conclusions with respect to the order of learning tasks, difficulties in reading the map, and readiness were drawn.

A Comparative Study of Language Attitude and Intelligence in Sixth-Grade Children from Low-Socioeconomic and Middle-Socioeconomic Levels

CAROLYN EMRICK MASSAD, Educational Testing Service

Research has shown that the results of intelligence measures may be affected by language differences and/or cultural deprivation. However, the need to determine the interrelationship between language aptitude and intelligence is apparent.

The major objectives of this study were to: (a) work toward a clearer definition of the term language aptitude; (b) better define the relationship between language aptitude and intelligence; and (c) clarify the role socioeconomic level has to play in determining this relationship.

The subjects were 132 sixth-grade, public school pupils, 93 of which were from a middle-socioeconomic level area and 39 of which were from a low-socioeconomic level area of northeastern Ohio. Distinction regarding socioeconomic level was made by using the Otis-Dudley-Duncan Socioeconomic Index. The general plan of the study involved the administration of two types of instruments: the MLAT (*Modern Language Aptitude Test*) to measure language aptitude; and the SCAT (*Cooperative School and College Ability Tests*) to measure intelligence. The experimental design basic to this study was a factor analytic design.

The factor structures from both the low- and the middle-socioeconomic level areas indicated that attributes labeled language aptitude and intelligence have a great deal in common. However, both the correlation matrixes and the factor structures showed that sixth-grade children from differing socioeconomic levels employ different processes in thinking about language. The children from the middle-socioeconomic level tend to approach all tasks involving the interrelations among meaning, sound, and symbols in much the same way, and in a way different from their approach to sentence structure. In contrast, the sixth-grade children in the low-socioeconomic level group seem to have separate approaches for structural relations, for sound-symbol-meaning relations, and also for recognition of symbols. For the low-socioeconomic level group, it appears that the more clearly sound and form of words are related in all aspects of the language arts program, the more consistent and the more broadly useful would be the results.

The data presented have led the investigator to believe that language aptitude is not a unified dimension of the cognitive processes (as is intelligence) which is independent of intelligence. However, the present research has shown that the differentiation of abilities contributing to language aptitude in sixth-grade children is distinct from that which past research has indicated for older children and adults.

The Characteristics of Advantaged and Disadvantaged Pupils' Teachers

ALBERT H. YEE, University of Wisconsin

In efforts to improve educational opportunities in the Great Cities, little empirical research has been applied to differences between the characteristics and methods of teachers teaching advantaged and disadvantaged pupils (Gage, 1965; Gordon, 1965). A study intended in part to overcome such lack, recently completed by Coleman, et al. (1966), has been a source of great debate. Some problems, according to critics (e.g., Nichols, 1966) are: Teacher-pupil interaction in classrooms is neglected; teachers' and pupils' characteristics are considered separately, and schools are considered the unit of analysis. The report recommended that for equality of educational opportunity, schools must provide "strong effects" independent of disadvantaged pupils' immediate social environment and "good teachers." What "strong effects" and what characterizes "good" teaching are glossed over in the voluminous work.

The present report (Yee, in press)provides ANOVA investigations of the attitudes of 212 intermediate-grade teachers (MTAI) and their pupils ("About My Teacher")—102 advantaged and 110 disadvantaged classes. Factor analyses helped provide a total of 5 teacher and 12 pupil measures. Pretests were administered about the second week of school, and posttests were administered toward the end of the first semester.

Significant 2-way ANOVA results indicate pupils' social-class background to be the most potent source of variance. Second most important contributor was teachers' sex. Grade level and teachers' years of experience were next in order. Significant interaction effects were found operating between pupils' social-class background and teachers' sex; also, between pupils' social-class background and grade level. One-way ANOVA results showed significant differences between four levels of pupils' ethnic and social-class characteristics.

In improving educational opportunities, the results indicate strong need to consider administrative and community factors, such as teacher selection, salaries, retention rates, etc., and personality and pedagogical characteristics of teachers, for effective placement with disadvantaged pupils.

3:00 to 3:55 P.M. **SESSION 7.6** **DIVISION D**

READING RESEARCH

CHAIRMAN:

Frank P. Greene, Syracuse University

PAPERS:

Identification and Measurement of Reading Skills of High-School Students

FREDERICK B. DAVIS, University of Pennsylvania

This study was designed to estimate the per cent of the nonchance variance of each of the most important skills in reading comprehension among mature readers that is unique in the set of skills measured.

Eight skills judged of greatest importance in reading comprehension were selected. Forty multiple-choice items to measure each skill were tried out in a large sample of twelfth-grade pupils in academic high schools. Each item was based on a separate passage.

A differential item analysis was performed in such a way as to obtain eight uninflated biserial correlation coefficients for each of the 320 items between passing or failing the item and scores on each of the eight skills. For each skill, 24 items were selected that had maximum correlations with the skill that they were intended to measure and minimum correlations with the seven skills that they were *not* intended to measure. Two equivalent forms of 12 items each were assembled to measure each skill. These tests were next administered to a new sample of twelfth-grade pupils; 988 pupils turned in complete answer sheets on all of the 16 forms. Intercorrelation matrixes of 50 scores for each pupil were obtained in the entire sample and in each of two random halves of it. Uniqueness analyses were then performed, cross-validated by items and, separately, by examinees.

Surprisingly large percentages of the nonchance variance of five of the reading skills were found to be unique in the set of eight skills. The two largest segments of unique variance were found in skill scores measuring memory for word meanings and drawing inferences from content. It is clear that reading comprehension among mature readers is not a unitary trait. Consequently, a workbook of specially designed learning exercises will be prepared to develop proficiency in the separate reading skills. The technique employed in this study should next be applied to comprehension skills among elementary-school and among junior high-school pupils.

◆ ◆ ◆

Poor Readers' Functioning on the WISC, Slosson Intelligence Test and Quick Test

CAMILLE HOUSTON and WAYNE OTTO, University of Wisconsin

The purpose of the study was to examine the potential usefulness of two short, easily administered individual intelligence tests—the Slosson

Intelligence Test (SIT) and the Quick Test (QT)—with poor readers in the elementary grades. The reading demands of most group intelligence tests make them generally inappropriate for use with pupils with reading disabilities; yet, the most respected alternatives, the Wechsler Intelligence Scale for Children (WISC) and the Stanford Binet, require much time and sophistication on the part of the examiner. Thus, there is a need for individual tests that can be administered and scored quickly and in a straightforward manner and yield reasonably valid scores with poor readers. In the present study the relative validities of the SIT and the QT were examined by comparison with the more comprehensive WISC.

The subjects were 56 elementary pupils who had been selected for enrollment in a summer reading program on the basis of (a) average or better intelligence according to test scores and/or teacher judgment, and (b) a marked gap between capacity and achievement in reading. WISC scores, derived from tests administered by qualified school psychologists, were available for all subjects. Trained examiners administered the SIT and the QT to each subject. Means, standard deviations and intercorrelations were computed for chronological age, SIT IQs, QT IQs, WISC Verbal Scale IQs, WISC Performance Scale IQs, and WISC Full Scale IQs; and discrepancy comparisons of WISC, SIT and QT IQs were made.

The data indicate that SIT and WISC (Full Scale) IQs ($r = .60$) are significantly more closely related than are QT (Total) and WISC (Full scale) IQs ($r = .49$) although scores from both of the brief tests were significantly correlated with WISC scores. The results are discussed in terms of test characteristics and implications are drawn in terms of purpose for testing.

◆ ◆ ◆

Canonical Analysis of Reading Readiness Measures and the Frostig Developmental Test of Visual Perception

FRED W. OHNMACHT and ARTHUR V. OLSON, University of Georgia

The present study sought to explore the communality among several measures of reading readiness and the Frostig Developmental Test of Visual Perception (FTVP). *Ss* consisted of 232 first grade children from three semi-rural schools in Northern Georgia.

All *Ss* were administered the Metropolitan Reading Readiness Test (MRRT), the Gates Reading Readiness Test (GRRT), and the FTVP during the first month of school. Three canonical correlation analyses were undertaken to determine the common functions underlying the three batteries. The FTVP was correlated with the MRRT and GRRT in turn. The MRRT was also correlated with the GRRT.

The analysis of the FTVP with the MRRT produced one significant canonical variate ($Rc = .802$, $p < .001$). This variate was defined by the

Spatial Relationships and Position in Space sub-tests of the FTVP on the one hand, and by the Copying and Matching sub-tests of the MRRT on the other. This finding replicates a factor identified as Perceptual Readiness in a prior study by one of the authors. The FTVP, GRRT analysis also yielded but one significant variate ($Rc = .732$, $p < .001$). The Spatial Relationships sub-test of the FTVP again was the major contributor to the FTVP readiness variate, whereas Picture Directions and Word Matching were the major determiners of the variates composition selected from the GRRT.

The GRRT, MRRT analysis produced two significant canonical variates. Interestingly enough, the first ($Rc = .827$, $p < .001$) variate was largely determined by the same variables which made the major contributions to each battery's significant canonical relationship to the FTVP. From these analyses it was concluded that the MRRT and GRRT possess one underlying common perceptual function reflecting some of the information yielded by the FTVP. Other studies suggest that the information unique to FTVP yields little that is useful in predicting reading achievement.

◆ ◆ ◆

An Analysis of the Relationship Between Conventional Reading Readiness Measures and Intellectual Functioning

NORINNE H. OLSON, 9th District Educational Services Center

The purposes of this study were to examine the degree of communality between selected reading readiness tests and to determine the predictive power of readiness tests and sex as variables contributing to intelligence test performance.

The following hypotheses were formulated:

1. Communality exists between reading readiness subtests such that the subtests of two commonly used reading readiness tests do not measure separate abilities.

2. A. Reading readiness test scores are not independent of intelligence test scores, such that a significant proportion of the variance of intelligence test scores can be predicted from reading readiness test scores. B. In the establishment of the prediction scheme, the sex of the respondent will be a differential predictor.

The subjects included in the population of this study were 294 children enrolled in fifteen first grade classrooms of two rural school systems in northeastern Georgia.

All subjects were tested using two reading readiness measures (*Metropolitan Readiness Test; Lee-Clark Reading Readiness Test)* and one group intelligence test (the *California Short-Form Test of Mental Maturity,* Level 0).

The data were subjected to two analyses to investigate the questions posed the two hypotheses.

A factor analysis of the nine reading readiness subtests indicated the presence of a large general factor with contributions from all nine variables. The eigenvalue for this one factor accounted for almost one-half of the proportion of variance of the nine variables in the analysis.

A multiple linear regression analysis was employed to test the readiness-intelligence relationship, using language, non-language, and total I.Q. scores as the criterion variables and readiness test scores, age, and sex as predictor variables.

F values for full models indicated that the predictor variables were significant predictors of the three intelligence areas. F values for the significance of the decrease when sex was deleted from the full model were not significant.

The results of these analyses appeared to indicate that administration of more than one reading readiness test is impractical since the tests seem to measure the same abilities.

3:00 to 3:55 P.M. **SESSION 7.7** **DIVISION B**

LANGUAGE ARTS

CHAIRMAN:

O. L. Davis, University of Texas

PAPERS:

Comparison of Three Methods of Beginning Reading Instruction: ITA, DMS, TO

EDWARD FRY, Rutgers, The State University

Three methods of beginning reading instruction were compared at the end of 1st, 2nd, and 3rd grade on the Stanford Achievement Test, the Gilmore Oral Reading Test, and other measures. The three methods were: (1) The Initial Teaching Alphabet (ITA) Mazurkiewicz and Tannyzer materials; (2) a traditional set of basal readers (TO) called the Sheldon Readers; and (3) a set of the Sheldon Readers marked with the Diacritical Marking System (DMS). At the end of all three grades, there were no differences on the Stanford or Gilmore tests. The ITA group had inferior spelling at the end of 1st grade but not at the end of 2nd or 3rd grade. The ITA group wrote longer stories and could read phonetically regular words in isolation better than the other two groups. Each group contained 7 classrooms in the 1st grade and were located in suburban schools with intelligence scores very near the U. S. norm. Analysis of variance did not indicate any method to give superior reading achievement for bright or dull students, or for girls or boys.

Seven new classrooms were started the second year with a new DMS set of materials. These children were tested on the same measures at the end of their 1st and 2nd grade, and compared with the original results. There were no significant differences between the new DMS group and the three preceding groups, though there was a tendency for the new DMS group to do better on some reading measures than the old DMS group.

A large number of correlations were computed at the end of each year. Using the Stanford paragraph sub-test as the criterion of reading achievement, we found the following:

a. There was little correlation with age.

b. There was little correlation with class size.

c. Students who could read before entering or were ahead at the end of 1st grade tended to stay ahead.

d. Having a good teacher in 1st grade was important.

e. Reading readiness tests were not good predictors.

Last, but not least, we found that special alphabets for beginning reading instructions had been tried in earlier centuries and abandoned.

A Study of the Effectiveness of a Structured and Sequenced Program of Vocabulary Study in Improving the Reading, Writing, and Listening Skills and the Verbal Abilities of High-School English Students

MILTON D. JACOBSON, University of Virginia; BETTY YARBOROUGH, Director of Language Arts, Chesapeake; ELIZABETH B. HANBURY, Supervisor of English, Chesapeake

No methods studies of vocabulary development at the secondary level have been reported during the last two decades. Yet high school students are required to have considerable sophistication in vocabulary development for successful performance on currently popular achievement and abilities tests.

The purposes of this study were: (1) to investigate the effect of a sequenced vocabulary program upon the reading, writing, listening, and spelling achievements of students in grades 10, 11 and 12: (2) to determine the effect of the program upon verbal ability and subsequently total scholastic ability; (3) to study the effect of the program upon the relationships among the verbal skills listed and between verbal achievement and ability; (4) to predict SAT scores by combinations of test scores.

Nine hundred experimental subjects (taught by nine high school English teachers) and three hundred controls were pre-tested and post-tested. Criterion measures included STEP and SCAT, *Brown-Carlsen Listening Test, Nelson-Denny Reading Test, Wide-Range Achievement Test* (spelling), and SAT. Variance and co-variance techniques were used to analyze the data.

The first year of this two-year study has been completed. Results of preliminary analysis of the data were encouraging. Significant differences were found which appear to have important implications for methods and instructional materials in teaching high school English, particularly as they concern vocabulary development.

◆ ◆ ◆

Efficacy of Specific Training in Listening—Phase 11

ALBERT M. DeSOUSA, Scotch Plains—Fanwood School System, Scotch Plains, New Jersey, and MILLY COWLES,* University of South Carolina

DeSousa and Cowles (1967) provided 20 days of special training in purposive listening to a group of 30 randomly selected seventh-grade subjects. An isolated-control group (30 subjects) received 20 days of special training in thematic literature study and a control group (30 subjects) received no special treatment. Subjects were similar in listening ability as measured by the *STEP: Listening Section* (Form 3A) and in intelligence as measured by the *Wechsler Intelligence Scale for Children* before the experiment began. Post-tests, *STEP: Listening Section* (Form 3B), were administered to all groups one week after the experimental treatment was terminated. Analysis of variance performed on differences between pre- and post-test scores produced an F of 9.41 ($p < .01$). The mean of the experimental group differed significantly when compared (Duncan's New Range Multiple Test) with the isolated-control group and control group's means. In brief, the experiment was successful in terms of the performance of the experimental group on the dependent variable, listening ability.

A major question raised by the experimenters was whether the gains shown by the experimental group in listening ability would be maintained after a substantial lapse in time. The purpose of the study reported herein was to measure listening ability of the three groups one year after the close of the original experiment.

Hypothesis tested was that the *STEP* scores would not be significantly different for the experimental, isolated-control, and control groups one year after the close of the purposive listening study.

All subjects were located and re-tested one year after the close of the first DeSousa-Cowles listening study. The *STEP: Listening Section* (Form A) was used to measure subjects' listening ability. A Two Factor Repeated Measures Analysis of Variance was performed on pre-, post-, and post-post-scores for the three groups.

As in the original experiment, significant differences were obtained among the three groups in listening ability. The analysis shows that the groups differed in listening ability, differed at various stages in the experiment, and differed when the interaction between the two (listening ability and stages) was compared.

The Newman-Keuls method for testing the difference between all possible pairs of means showed, among other things, that the experimental group's mean differed significantly from isolated-control and control groups' means after one year had lapsed.

The experimental group maintained the gains that were shown one year ago. Clearly, the listening treatment did not dissappear with the passage of time.

◆ ◆ ◆

A Comparison of the Effectiveness of Five Exotic Methods of Remedial Reading Instruction

JOSEPH C. JOHNSON, II, Duke University, and MILTON D. JACOBSON, University of Virginia

This study examined a number of 8, 9 and 10 year old remedial readers in an effort to determine which grouping of materials is superior in teaching them in a clinical situation. Specifically this study was designed to determine if the relationship between number of years in school and recalled comprehension response, and the relationship between initial levels of instruction and recalled comprehension response, are significantly effected by any of the following specified materials and techniques: linguistic, programmed, phonetic, kinesthetic, or machine, when considered in conjunction with the dependent variables of intelligence, chronological age, sex, socio-economic status, reading achievement, grade level, race, and grade placement level.

The procedures utilized included: (1) The WISC and the Metropolitan Reading Achievement Test were administered to approximately 1000 children of ages 8, 9, and 10; (2) A sample of 86 children representative of national remedial reading norms in intelligence and reading ability were selected from 30 elementary schools in Virginia; (3) the selected children were randomly assigned to teachers and specific materials grouping technique; (4) these children received instruction in these specified methods and techniques for a period of 8 weeks; and (5) computerized co-variance techniques were utilized to control variables found to effect relationships between materials, groupings and techniques and recalled comprehension response.

Significant differences (.05 level) were found between methods (materials and techniques). There were significant interactions between grade placement level and chronological age, between grade placement level and I.Q. and between method and grade placement level.

3:00 to 3:55 P.M. SESSION 7.8 DIVISION E

COMPENSATORY AND VOCATIONAL EDUCATION

CHAIRMAN:

Raymond C. Hummel, University of Pittsburgh

PAPERS:

A Study of Attitudes Towards the Urban Community Among Disadvantaged Male College Students

ROSCOE C. BROWN, JR., HARRY PLOSKI, DOROTHY LANDER, New York University

An instrument known as the Urban Perception Test was used to identify the attitudes of 27 students in Project APEX, a college program at New York University for youth from disadvantaged backgrounds, towards minority group life, slum living, and the opportunity structure in the urban community. The APEX students' attitudes were compared to the attitudes held by samples of lower class Negro and white youth and by middle class Negro and white youth in the New York Metropolitan area. A series of eleven photographs was chosen to represent certain aspects of urban life and a set of questions to follow each photograph was developed and pretested.

The results revealed statistically significant differences between the APEX students and the other samples tested in their view of life in the urban community. The APEX students were ambivalent about the educational system. Some felt that education would not lead to a better life, while others saw education as a way to a better job. The APEX students saw the white middle class family as having fewer problems than those of the Negro family whether lower class or middle class; however, the APEX students differed from the lower class Negro sample in their view of the opportunity structure of the society. The APEX students more frequently viewed the urban slum as being caused by poverty rather than prejudice, and felt that it is possible for them to improve their status through a better job or education.

◆ ◆ ◆

Evaluating the Effectiveness of a Compensatory Educational Program on the Higher Education Level for a Culturally Deprived Group

ARYE PERLBERG, University of Illinois and YAEL RAM, Israel Institute of Technology.

In 1963, the Technion, the Hebrew University, the Israel Defense Army, and the Ministry of Education established a one-year, pre-university preparatory program for soldiers of Oriental origin, in order to increase their numbers in institutes of higher learning. The ultimate purpose was to increase the number of professionals and highly educated political, culcultural, and economical elite of Oriental origin.

Five groups (396 students) participated in the program; most continued their studies. The first group has already been graduated. The purpose of this study is to evaluate the program effectiveness as reflected in academic and non-academic achievements. Attention will be given to the role of cooperating organizations (Army) in this social experiment.

Five areas are investigated: (1) *Documentation*—a sequential analytical description of the program, its problems, solutions, and a demographic socio-economic and intellectual profile. (2) *Academic* evaluation of selection instruments and procedures. Participants' academic growth is analyzed and compared to their own achievements in various stages and to control groups. (3) *Dropouts* follow-up studies, reasons, alternative routes, and variables differentiating between dropouts and graduating participants. (4) *Social* evaluation includes social acceptance, involvement, and status among other students. (5) *Attitude* studies of various Israeli groups towards this program.

Selected findings from interim reports are: subjects are highly selected group among Orientals. Lack of formal qualifications and motivation reduces number of candidates. The preparatory course helped most to succeed in competitive entrance examination (90-95%) as compared with the whole student population (35-45%). Preparatory program was not sufficient, and learning difficulties required additional tutoring of freshmen and sophomores. Over all, achievements similar to general Technion average. Participants were active in student body. Program duration should be extended and curriculum expanded for more fundamental enrichment and compensation. Participation in program during last year of Army service facilitates solutions to organizatonal problems and contributes to Army's positive image.

◆ ◆ ◆

Rioters' Attitudes Toward Education: Interviews with Participants

FRANK P. BESAG, State University of New York at Buffalo

Immediately following the racial disturbances in Buffalo during the week of June 26 through July 1, 1967 a study of the events and causes of the disturbance was begun. The distinguishing characteristic of this study was that rather than go to the "leaders" and ask them for their opinions, 138 youths and adults who had viewed incidents or who had actually participated in the incidents were interviewed. Most of the interviewees were Negroes who lived in the disturbance area. Their ages ranged from ten to seventy with the greatest number being between sixteen and twenty. There were also some comments from lower socio-economic Caucasians. All of these were adults. The study discusses the interviewees' attitudes, not facts, about the school.

One of the concerns raised by many of the interviewees (forty-three) was that they felt that their education had been inferior and that this caused a great deal of frustration and anger within the Negro community. Specifically, the concerns expressed were: 1. the most general comment made was that the teachers and administrative staff had no respect for them or their "culture" and therefore, could show little respect in turn; 2. that what they learned in school had little or no relationship to what they needed to know in their daily lives; 3. that youths were counseled out of academic programs; 4. that youths were treated summarily and had no appeal in matters of discipline; 5. high teacher turnover; 6. that the vocational education did not prepare them for future employment.

The paper will use quotations from the actual interviews plus an analysis of the data.

◆ ◆ ◆

An Interdisciplinary Approach to Vocational Education

RAYMOND J. AGAN, Kansas State University

The basic thesis tested in a program of research conducted over the past two and one-half years in the public secondary school of Paola, Kansas, was that the matter of occupational information, selection, and preparation no longer falls within the province of the several vocational fields, but that these problems are the concern of persons representing a broad spectrum of disciplines. Evidence of changes from the discipline oriented point of view was presented in the 1943 and 1965 yearbooks of the National Society for the Study of Education as well as the 1963 Act for Vocational Education.

A team of teachers under the leadership of a program director functioned as a coordinated group to present a program of occupational exploration and study of the common elements of the vocational disciplines to a group of eleventh year students. The team, who were also teachers of traditional classes in the secondary school, supervised and counseled with the students as they observed occupations of their choice in the business concerns of the community.

In a process of self-analysis of aptitudes, interests, and attitudes toward work, the students selected a field for occupational experience to be carried out in the twelfth year under the supervision of the team of teachers and local cooperating business men for a minimum of fifteen hours weekly.

Evaluative techniques used to appraise this system of interdisciplinary vocational education included an intensive community survey, follow-up surveys of the students, a comparative use of the GATB test, and opinionaires from the students, parents, cooperating business men, and the teachers.

Conclusions included a positive correlation toward interdisciplinary vocational education, the endorsement of the rationale and plan as developed from the research by the Mid-Continent Regional Educational Laboratory and the TEPS Commission of NEA as a Year of the Non-Conference Demonstration Center and and is open to visitors.

3:00 to 3:55 P.M. **SESSION 7.9** **DIVISION C**

GROUP CHARACTERISTICS

CHAIRMAN:

Dale E. Mattson, Association of American Medical Colleges

PAPERS:

Sociometric Choice: A Study in Pupillary Response

ROBERT H. KOFF, Stanford University and THOMAS H. HAWKES, InterAmerican Education Center

Pupillary dilatation has recently been reported as a physiological measure of degree of attention paid to environmental stimuli (Hess and Polt, 1960, 1964). Hess and Polt's research has operationally defined one variable and has provided data that allows us to speculate about several others. This variable is interest-disinterest. The purpose of this investigation was to examine the relationship between sociometric choice patterns and pupillary response. Twenty-seven sixth grade *S*s completed a sociometric questionnaire in which they indicated 3 friendship and 3 nonfriendship choices. Black and white photographs of each *S* were made and served as the experimental stimuli. The stimuli were arranged and shown to each *S* according to their responses to the sociometric questionnaire. Differences in pupil dilatation patterns were found among *S*s when viewing stimuli depicting friendship choices as opposed to nonfriendship choices. There was significantly ($p<.05$) more pupillary dilatation, on the average, to pictures of friends than there was to pictures of nonfriends. Results were discussed in terms of a) theoretical issues relevant to sociometric choice patterns; and b) pupillary response patterns and their relationship to inquiry in the social sciences.

◆ ◆ ◆

Creativity and Conformity as Functions of Group Size and Instructions

THOMAS ANDRE and HARRY SCHUMER, University of Massachusetts

The study was concerned with determining the effects of discussion group size and test instructions on individual creativity and conformity. Discussion groups of sizes 3 and 6 were compared with no-discussion individuals under two types of directions, creative instructions, and quantity instructions. The criteria of creativity were the Word Association and Uses for Things tests developed by Getzels and Jackson (1963). A measure of conformity was obtained by deriving deviation scores on each individual from mean score of his group. The results indicated that both group sizes performed more creatively than did individuals, but that no differences existed between 3 and 6 groups. Ss who received creative instructions scored higher than did Ss who received quantity instructions; but differences appeared significant only for groups 3 and 6.

The conformity data indicated that 3 man groups conformed most, 6 man groups conformed second most, and individuals least. The results were discussed in conjunction with results from previous studies. It was concluded that group size and instructions have important effects on the creative behavior of groups. Other variables such as type of task, and the type of subject, may modify the effects of group size and task instructions. The relationship between conformity and creativity was seen as more complex than had heretofore been considered. The effect of conformity on creativity seemed to depend on other relevant variables in the group situation. Subsequent researchers in this area should bear these conclusions in mind and plan research with a view towards controlling the relevant factors of group behavior.

◆ ◆ ◆

Similarity of the Group's Perception of Its Interrelational Structure With Group Member's Personality

R. ROBERT RENTZ, EMORY B. FEARS, WILLIAM F. WHITE,
University of Georgia

Application of Traditional Techniques of sociometry to measurement of the interrelational structure of classroom groups has enjoyed some degree of popularity among educators. Rentz and Olson (1967) have criticized these methods as being too situationally specific. They advanced the notion of using semantic differential (SD) data to determine group structure, hypothesizing that these data would have greater situational generalizability. Using the methods they describe, a group member is rated on a set of SD scales by his fellow members, enabling each group member to be located as a point in a multidimensional semantic space. The group structure can be described by the relational pattern of these points. The pattern thus represents the group's perception of its members based on a perception of affective cues. This study sought to determine the relationship between the personality of group members and their position in the group structure.

High school students in three classes of Remedial English comprised the 3 groups ($N_1 = 16, N_2 = 17, N_3 = 16$). Each group member rated all members of his class on a 12 scale *SD*. A factor analysis of these ratings defined a three dimensional space, with the 3 factors corresponding to Osgood's primary three semantic factors of Evaluation, Potency, and Activity. Mean factor scores for each member determined that member's group location. Ss were also administered the *High School Personality Questionnaire* (14 factors). Finally, the 3 *SD* scores were correlated with the 14 *HSPQ* scores by the canonical method.

The first canonical correlation (.793) was statistically significant ($p < .02$). The assocated canonical vector contained major weights for the

HSPQ factors: E, G, I, $-Q_2$, B and D; as well as a strong negative weight for the *SD* Evaluation factor. This correlation demonstrated the substantial relationship between the group's perception of its members and those members' personality. Apparently, the groups perceived the independent, perservering, somewhat bright individual, who was sensitive to group needs, as inadequate, undesirable, and of less worth. Some degree of defense against competent group leadership seems to be evident in the insecurity of these remedial students.

◆ ◆ ◆

Influence of Method of Questioning Upon Children's Responses to Humorous Situations in Literature

DIANNE L. MONSON, University of Washington

The study was concerned with a comparison of two methods for eliciting children's responses to humor in literature. Fifth grade children read seven excerpts of humorous literature, reproduced from the books by means of xerox and ditto. The five broad categories of humor considered in selection of the excerpts were laughter at characters, surprise, the impossible, use of words, and the ridiculous situation. Excerpts were chosen from the following books: HENRY HUGGINS, CHARLOTTE'S WEB, PIPPI LONGSTOCKING, OWLS IN THE FAMILY, and THE CRICKET IN TIMES SQUARE.

An instrument for eliciting responses to the selections was developed in four forms. One form asked that children judge whether the excerpt was humorous and, if the response was affirmative, write an explanation of why the excerpt was funny. The other forms also asked children to judge whether the excerpt was humorous, but asked children to indicate the source of humor by responding to multiple choice or true-false questions about the excerpt. Questions were constructed from written responses made earlier by a pilot group.

Children were grouped for analysis of data on the basis of sex, intelligence, socioeconomic level, and reading ability, as well as response forms. Responses were analyzed in order to determine the influence of the writing task upon the judgment of an excerpt as humorous or not humorous. The three-way chi-square statistic was used for analysis of data.

Findings were as follows:

1. Differences in responses made by children in structured and unstructured treatment groups were greater for boys than for girls. Boys more frequently judged selections humorous when they were presented in a structured situation.

2. Children in low socioeconomic groups more often judged selections humorous when they were presented in a structured situation.

SESSION 8.3

9:00 to 9:55 A.M. **DIVISION C**

MATHEMATICS INSTRUCTION

CHAIRMAN:

Hulda Grobman, New York University

PAPERS:

Generality of Heuristics as an Instructional Variable

JAMES W. WILSON, School Mathematics Study Group, Stanford

This study investigated the problem solving performance of *Ss* on learning tasks and transfer tasks following instruction in which the level of generality of the problem solving heuristics taught was an independent variable. It was predicted that *Ss* taught more general heuristics will tend to perform better on transfer tasks, whereas *Ss* taught task specific heuristics will tend to perform better on the training task measures. Three levels of generality in problem solving heuristics were used: task specific heuristics, means-ends heuristics, and planning. The theoretical framework was derived from the research on computer simulation of complex human problem solving and the writings of Polya.

144 *Ss* who had completed a course in high school plane geometry were trained in two tasks with self-instructional booklets. Independent variables were level of heuristics used in training on each task, and order in which the two training tasks were taken. A 3 x 3 x 2 factorial design was used, and *Ss* were randomly assigned to one of the treatment combinations. The dependent variables were time required to solve a set of problems in each of the training tasks, two transfer tasks similar to the training tasks in format, and three transfer tasks unlike the training tasks in format.

The first part of the general hypothesis, that task specific heuristics would facilitate problem solving on the *training task* better than general heuristics, was not supported. For transfer task variables, two types of interaction were found. The first can be called "complementary heuristics." Problem solving performance is enhanced by the combination of different types of heuristics on the two training tasks. The second can be called "practice of transfer." This is an interaction of level of heuristics with order of training.

◆ ◆ ◆

Language of Exposition and its Influence on Mathematics Learning

S. M. AVITAL and D. SUGERMAN, University of Toronto

The problem of the influence of the language of exposition on mathematics learning was raised by Henderson (1962). A study by Avital (1967) showed that learning the structure of the Klein group (when given under three different treatments: verbal expositive, mathematical symbolic, and patterning) is most effective when the student is exposed to the material under a patterning treatment in which the mathematical meaningfulness of the structure is sacrificed, and the structure is taught as a symmetrical array. In the same study a significant grade effect was obtained, with students in higher grades performing better than students in lower grades. This effect was interesting because all students were at an age when the level of formal thinking is reached and the material learned was independent of the school curriculum. The criterion variable for all three treatments was a twelve-item retention test prepared by the investigator.

As a result of this study the following questions were asked:

1. Will the superiority of the patterning exposition be maintained when a less symmetrical structure is used?

2. Will the same differentiation in performance be maintained on a criterion for which the need of transfer mechanisms could be postulated?

3. Will the perennial inferiority of females in performance of mathematical tasks (Kagan, 1963) be maintained in a task independent of general school learning, and in which the competitive nature of performance on school tasks is not emphasized?

The design of the study was a 3 x 3 x 2 factorial (Treatment x Grade x Sex). The analysis of the results of the total test showed significant main effects due to Treatment ($p<0.001$) with the symbolic method being the most effective. Also significant was the main effect Grade ($p<0.001$) with the higher Grades performing better. The main effect Sex was not significant ($F= 0.23$). The Interaction Grade x Treatment was not significant ($F = 1.75$). Essentially similar results were obtained separately for each subtest.

The above results show that when the natural structure is less symmetrical, pattern storing leads to less effective storage than meaningful relationships (such as inverse element, neutral element, etc.). However, a verbal description of a meaningful structure does not improve learning. Apparently the tendency of some teachers to use verbal expositive description of a structure may be superfluous, and perhaps even harmful. The nonsignificant main effect due to sex may point to the fact that sex difference in classroom performance in mathematics is mainly culturally induced.

The significant effect due to Grade supports the results obtained in the previous study that continuous involvement in learning mathematics transfers positively to learning of new structures.

◆ ◆ ◆

Teaching Mathematics by Exposition

JOSEPH M. SCANDURA and ROBERT G. McGEE*, University of Pennsylvania

The general purpose of this research was to identify factors affecting the learning of hierarchically arranged expository material, and to help determine the relative strengths of these factors. Four experiments were conducted. An artificial, hierarchically ordered subject matter was used in the first three experiments. The materials used to denote the ideas introduced at the lower (prerequisite) levels were, in each case, also used in describing the high order (i. e., higher level) material.

The purpose of Experiment One was to determine the effects of practice on prerequisite tasks and on problems defined at the criterion level on the solving of additional criterion-type problems and transfer problems. Experiment Two was designed to determine whether information about the nature of problems is necessarily reflected in problem solving performance. The third experiment was concerned with the effects of, and interactions between: (1) amount of prerequisite knowledge, (2) order of presentation of the prerequisite and criterion materials, and (3) prerequisite practice. The fourth experiment was conducted to determine the feasibility of extrapolating to mathematics from the previous results.

Four major conclusions were drawn from this research.

1. Practice on prerequisite tasks significantly improves the learning of higher order material, as judged by criterion test performance, when the prerequisite terminology is used to describe this higher order material.

2. Knowledge had by a learner affects future learning ONLY when this knowledge is prerequisite to the to-be-learned material.

3. The possibility that prerequisite learning simply amounts to learning part of the criterion task, or that a substantial amount of criterion material can be retained and then correctly interpreted when the prerequisite meanings later become available, is untenable.

4. Simple exposure to prerequisite information is often not sufficient to insure later learning. Furthermore, spending more time on higher order material, and having an opportunity to practice on related criterion problems, cannot normally be excepted to overcome prerequisite inadequacies.

◆ ◆ ◆

Are the Basic Skills Being Neglected in the Modern Mathematics Program?

VIDYA BHUSHAN and JAMES JEFFRYES, University of Hawaii

The purpose of this study is to compare students in a modern mathematics program with those who study the traditional mathematics with respect to their competency in basic computational skills.

Two groups of high school seniors were taught trigonometry during the fall semester of 1966, one with a UICSM approach and the other with a traditional approach. The Cooperative Test in Plane Trigonometry was administered to each group when they had completed the prescribed units of study in trigonometry to determine their achievement level. With a view for equating the groups in terms of intelligence, the California Test of Mental Maturity was given in the beginning of the course.

At the end of the experiment complete data were available for 59 students—24 in the UICSM group and 35 in the traditional group. Since the students were not assigned randomly to the two groups when the experiment began, two-way analysis of covariance was used to analyze the data, using IQ as a covariate. The UICSM group did significantly better than the traditional group on the trigonometry achievement test, $F(1,55) = 16.75$, $p < .01$. Also boys did significantly better than girls, $F(1, 55) = 5.52$, $p < .05$. Interaction between methods and sex was not significant, $F(1,55) = 1.78$, $p < .05$.

The results of this experiment indicate that students in a modern mathematics program become more competent at applying the basic skills to problem solving than students in a traditional mathematics program. This study supports the recent International Comparison of Math Achievement Study conducted by the International Project for the Evaluation of Education Achievement, in which students who had "new math' 'appeared to surpass those following traditional methods and boys tended to do better than girls.

SESSION 8.6

9:00 to 9:55 A.M. **DIVISION C**

PROGRAMED INSTRUCTION II

CHAIRMAN:

Philip W. Tiemann, University of Illinois at Chicago Circle

PAPERS:

The Effectiveness of Teacher Involvement in A Programmed Social Studies Sequence

FRANK L. RYAN, University of California, Riverside

Relative effects in terms of student achievement were determined for four instructional groups: teacher (P-T) or teacher aide (P-TA) involvement in a programmed instructional sequence; student working of programmed materials without teacher involvement (P); and no instruction (C). The N of 88 fourthgrade students was randomly assigned to the groups. The instructional model followed by both the teacher and the teacher aide is outlined herewith:

1. Teacher reads frame (sentences) from the programmed materials "using" the blanks:

2. Teacher allows student some time to write an answer;

3. Teacher randomly calls on student to supply the word (s) for the blanks;

4. Teacher indicates to the student if the answer given is correct or incorrect;

5. Teacher then rereads the frame and verbally inserts the correct answer, simultaneously holding up a card with the correct response on it;

6. Teacher goes to next frame of the programed materials and repeats the above sequence.

Neither the teacher nor the teacher aide was allowed to add any background information, new content, or further insight to the programmed materials on Japan's geography. The programs were identical for the P-T, P-TA, and P groups with one instructional program being used each day for five consecutive days. The criterion instrument was a 45-item multiple-choice test with a reliability of .89 (Kuder-Richardson, Formula 20) and covered specific facts, generalizations, and applications. Analysis of variance was used in treating the data. Posttest results indicated a superiority of both the P-TA and P-T groups over C ($p < .001$). On the retention test, the following contrasts were significant: P-T over C ($p < .005$), P-TA over C ($p < .05$), and P-T over P ($p < .05$). The results also indicated that (1) both the high reader (above the reading achievement median) and the low reader benefited from having a teacher or teacher aide involved in the programmed sequence (P-T, P-TA); and (2) criterion test gains took place for the low reader in the P-TA and P-T groups, but not for the low reader in the P group. The findings suggest that either teacher or teacher aide involvement in a programmed instructional sequence can enhance student learnings.

◆ ◆ ◆

Programmed Instruction as an Adjunct to a Course in Adolescent Psychology

LUTHER A. MARSH, Abilene Christian College, and JOHN PIERCE-JONES, The University of Texas at Austin

The purpose of this study was to investigate the differential effectiveness of programmed instruction as an adjunct to conventional university instruction. Learning was inferred from scores made by students enrolled in eight sections of a course in adolescent psychology taught by five psychologists, on a 100-item multiple-choice test, typical end-of-course examination. A 4 x 2 x 2 x 2 balanced factorial analysis of variance design was used to test the main effects of the independent variables (programmed instruction by two types of simple "teaching machines," by workbook, and control subjects receiving no programmed instruction; student determined laboratory schedule or instructor assigned schedule; traditional textbook or edited book of selected papers; the criterion multiple-choice final examination given as a pretest, or an essay pretest eliciting beliefs and opinions) on learning. Provisions were also made to control "Hawthorne effects," sex, and instructor variables.

Experimental subjects who received programmed instruction scored significantly higher on the criterion test than control subjects, but no significant differences between learning by teaching machines or by programmed workbook were observed. Subjects who received the criterion test as a pretest scored significantly higher on it when it was administered as part of the final course examination than did subjects who wrote about opinions or beliefs as their initial cognitive organizer or "set to learn." There were no significant differences in learning between subjects who chose their laboratory schedules, and those for whom schedules were assigned. Students who were assigned a traditional, integrated textbook "learned" significantly more than those who were assigned an edited book of readings. No significant sex differences were observed in learning and no significantly higher achievement was made by students of any of the five participating professors. A prediction of learning equation, using nine personal and cognitive variables for the 170 experimental subjects, yielded an R of .80.

The Effects of Subject-Matter and Individual Difference Variables on Learning From Scrambled Versus Ordered Instructional Programs

KENNETH H. WODTKE, BOBBY R. BROWN*, HAROLD R. SANDS, and PATRICIA FREDERICKS, The Pennsylvania State University

The results of several previous studies tended to support the somewhat surprising conclusion that scrambling an instructional program has little or no detrimental effect on learning (Roe, Case, and Roe, 1962; Levin and Baker, 1963; Payne, Krathwohl, and Gorden, 1967). The present authors hypothesized that the effects of instructional sequencing should depend on the conceptual characteristics of the subject-matter (following Gagnè, 1962) and individual differences among the learners (following Stolurow, 1964). The present investigators employed two subject-matters in two experiments designed to represent the extremes of a subject matter with a conceptual hierarchy (modern mathematics) and a subject matter consisting of relatively discrete facts (anatomy of the ear). Significant increases in within-program performance measures (errors and instructional time) in the scrambled sequence condition confirmed the expectation that the modern mathematics program did in fact contain a conceptual hierarchy. Lack of significant differences on within-program performance measures for scrambled and ordered versions supported the contention that the anatomy program lacked a conceptual hierarchy, and consisted primarily of a set of discrete facts. Randomly scrambled and ordered versions of the programs were presented to college *S*s via an IBM 1410 computer-assisted instructional system to test the hypotheses that (a) scrambling the sequence of instruction would have a detrimental effect on learning a subject matter containing a conceptual hierarchy as opposed to a program containing a relatively discrete set of facts; and (b) that the scrambled sequence would be more detrimental to the learning of low aptitude *S*s than high aptitude *S*s.

The results were interpreted in terms of the organizational strategies employed by college student learners. College *S*s given a scrambled sequence of instruction apparently are able to reorganize the material on their own, and are eventually able to achieve adequate mastery of the concepts, taking more time and making more errors in the process.

◆ ◆ ◆

Some Effects of Unit Structure on Achievement and Transfer

JEFF A. PYATTE, University of Virginia

Will there be significantly greater achievement and significantly more specific transfer, when programmed concepts in measurement are presented in a structured sequence as compared with the same material in a random sequence? Can the effectiveness of the two modes of presentation be measured? Can a test be devised to verify that a defined structure actually exists?

To begin to answer these questions, 199 fourth, fifth, and sixth graders were selected from schools which were participating in a larger study. A measure of basic arithmetic ability was made, and the students were separated,by grade, into three basic ability levels. The students were then randomly assigned a mode of unit, and they took the assigned unit as one unit in a programmed science course. Achievement tests were administered at previously defined points in the units, those points in the structured mode having the defined order, and those in the unstructured mode having a random order. The final achievement test was considered the criterion measure for achievement, and included transfer items from which a criterion transfer score was obtained.

The analysis of data was done using a mixed-model factorial design for analysis of variance to test the assumptions of equal achievement means and of equal transfer means. A regression analysis was used to give a measure of effectiveness and to decide the question of structure.

Mode of program was found significantly to affect neither achievement nor transfer. Grade level and basic ability level were found significantly to affect both achievement and transfer. Although difficult to separate, basic ability was probably the more important, the student with high basic ability having the higher achievement. For achievement, there was a significant interaction between mode of program and basic ability. Students of high basic ability achieved higher in the structured unit. By the test used, the units were found to be effective. By the test used, the structured unit was found to be structured.

SESSION 8.11

9:00 to 9:55 A.M. **DIVISION C**

CONTRIBUTED PAPERS II

CHAIRMAN:
David Crispin, Indiana State University

PAPERS:

Differential Patient Response to Instruction, Counseling, and Dental Treatment

DANIEL E. LUPTON, The University of Illinois at the Medical Center, Chicago.

A comparative pre-post study of the effectiveness of three approaches (dentistry, counseling, instruction) to the treatment of an alleged psychosomatic illness, Temporomandibular Joint Dysfunction, involved three groups totalling 60 chronic patients. The criterion variables were symptom reduction, increased knowledge of the disorder, and psychological adjustment. The principal motive prompting the experiment was that the treatment of TMJ dysfunction offered an opportunity to study three forms of important human activity in a single setting. Because of the complex interaction of the factors apparently involved in the disorder, it was possible to set up, under controlled conditions, two distinctive sets of educational experiences—an affective set, counseling; and a cognitive set, instruction; and to compare them with the effects of a non-educational set of experiences, a course of dental treatment. The results of the experiment viewed in the context of previous research indicate that dental treatment combined with an ancillary program involving personal interaction whether in an educational program of instruction or counseling is superior to dental treatment alone in the alleviation of painful physical symptoms of TMJ dysfunction; that a program of instruction and dental treatment is superior to dental treatment alone or combined with counseling in affecting the learning of information about the disorder; and that a program of counseling and dental treatment is superior to either dentistry alone or combined with instruction in decreasing a TMJ dysfunction patient's negative self-referrant expressions. This investigation, in demonstrating the effectiveness of the dentistry and education programs, underscores sharply the effects of personal relations and interaction upon the relief of painful symptoms; the effects of instruction upon learning; and of counseling upon negative feelings about oneself.

The Effect of Group Counseling on the Behavior of Educable Mentally Retarded Boys

PHILIP H. MANN, University of Miami, and MILTON D. JACOBSON and JAMES D. BEABER, University of Virginia

Emphasis placed on intellectual factors by many educators has resulted in only parenthetical interest in the personal and social adjustment of the retarded child, though this adjustment is important and may be beneficial to learning.

In most schools, counseling services are limited to normal children and are conspicuously lacking for special children. This is a questionable practice.

This study attempted, with group counseling procedures, to change self concepts of educable mentally handicapped boys wiith behavior problems. It also investigated the effects of counseling on the academic achievement, anxiety, deportment, and intelligence of these boys.

36 boys with ages ranging from 9-13 and I.Q.'s ranging from 56-80 were divided into control and experimental groups which were matched on chronological age, intelligence, and socio-economic status.

The experimental group received 12 one-hour counseling sessions from the principal investigator in a structured but supportive environment.

The control group spent equal time with the same counselor in a non-counseling, but structured, library-study situation.

Pre and post measurements were given and were subjected to analysis of variance and covariance.

The experimental group exceeded the control group in self concept improvement, anxiety reduction, academic achievement, and deportment (.05 level of significance).

This study, though essentially exploratory, demonstrates the benefits which educable mentally handicapped boys can obtain from group counseling and its effect on learning.

◆ ◆ ◆

An Investigation Into the Capabilities of the Generalized Academic Simulation Program (GASP) to Construct Four Kinds of Time Tables For Secondary School Use

JAMES B. KENNEY, University of Georgia

The generation of master schedules of classes has, until the emergence of the GASP, largely been a manual operation on the part of the principal of a secondary school. The purpose of the research was to determine levels of operating efficiency of the GASP when constructing four kinds of secondary school time tables.

Efficiency and capabilities of the GASP were investigated through four questions to which answers were ought. These were:

1. Can criteria other than those specified in the GASP be identified and defined that will improve the scheduling routine so as to result in fewer course request conflicts for the "Type I" high school schedule? (Type I schedules are those time tables where students meet six consecutive periods with little or no grouping of students by ability levels.)

2. Can an optimum program be developed for the (a) master schedule of classes, and (b) the individual student's schedule based on a heuristic approach? Presently, GASP constructs the master schedule through the heuristic approach, but *students* are sectioned into courses requested by an algorithmic technique. The investigator proposed to allow the computer to consider *both* the master schedule generation *and* student sectioning through the "power" approach. To allow the computer to explore all possible combinations of courses as opposed to placing a restriction to add or subtract certain course sequences for N trials as prescribed by the algorithmic method was proposed for this phase of the study.

3. Can the area of major teaching competence be improved through computer assignment of staff to sections of courses to be taught based on a numeric coding of first, second, and third areas of competence? The hypothesis being, that computerized assignment will maximize first areas of competence without having to resort to second and third areas to fill courses.

4. Can a GASP generated master schedule produce a more efficient schedule of classes in terms of: (a) number of student course request conflicts; (b) per cent of space utilized; (c) per cent of instructor time utilized within major area of competence; and (d) cost in terms of dollars spent?

(a) through (d), question 4 above, were compared to SHARE and NEEDS output using identical data bases.

Four secondary schools were selected for investigation. The four schools had all previously been computer scheduled using the SHARE and NEEDS programs. Students N's were 709, 1259, 1152, and 1400. Each school had uniquely different time table requirements. These requirements wer named "Type I, Type II, Type III, and Modular with released time." For each of the four questions heretofore posed, answers were determined and will be presented.

There were several critical problem areas identified; GASP documentation, coding incompatibility, maintaining integrity of classes tied across semesters by ability levels, providing "free time" through floating time modules, etc., all are discussed in the paper. Final run statistics for the four types of time tables are also presented (final runs for the four time tables yielded a minimum of 95% student course requests satisfied), together with a discussion concerned with manpower needs, cost in terms of dollars and time, and adequacy of satisfying the philosophical academic requirements of the secondary schools used as a data base.

SESSION 8.12

9:00 to 9:55 A.M. **DIVISION A**

CHARACTERISTICS AND BEHAVIORS OF SCHOOL PRINCIPALS

CHAIRMAN:

Joan R. Egner, Cornell University

PAPERS:

The Socialization of School Principals: Teaching and the Principalship

RONALD E. BLOOD, University of California, Santa Barbara

The problem undertaken in this study was to identify and describe elements from which theory could be designed to explain the functions and dysfunctions of teaching experience as preparation for practice in administrative roles within the school organization. The study specifically focused upon teaching experience as preparation for the elementary school principalship as a common instance of such role change. Consideration of the effect of teaching experience on performance in administrative roles in the school organization has centered most typically only upon length of experience rather than upon the nature of the experience itself. The problem area has been one virtually unexplored by researchers, but is of prime importance, as teaching experience is quite commonly required as a prerequisite for entry into non-teaching roles, such as the principalship, psychologist, counsellor, etc. The design of the study was exploratory, structured as an open theoretical scheme meant to discover variables in the socialization process rather than to test variables. Teaching experience was regarded as part of a social process, in particular the process of becoming a school principal. Recently appointed school principals were interviewed. The writer then attempted to evolve an explanation of the interview content through the development of theory from which testable hypotheses could be produced. It must suffice here to indicate that several hypotheses relate to predictions about the effect of alienation occurring in the process of anticipatory socialization. Others relate to the effects of early and late initiation of the process, that is, differences occurring in the practice of administration related to the earliness or lateness of the decision to become an administrator. A third set of hypotheses relate to the specific manner in which candidates for administrative positions gain access to the knowledge and perspectives of the administrative work world.

◆ ◆ ◆

Preferences of Principals for Instrumental and Expressive Characteristics of Teachers Related to System Type

EDWIN M. BRIDGES, Midwest Administration Center, The University of Chicago

This study was designed to identify the teacher characteristics to which elementary principals in middle-class and working-class type school systems attach the greatest importance when predicting the success of a beginning teacher in their system. Two classes of teacher characteristics were included in the investigation—instrumental and expressive. For purposes of this study, instrumental characteristics referred to those characteristics of the teacher that are concerned primarily with how she performs the task of imparting subject matter and building the intellectual skills of students. Expressive characteristics are those that are concerned with the kind of person the teacher is and how she relates *affectively* to students.

In this study the investigator manipulated the information which elementary principals were given on the instrumental and expressive characteristics of two teachers. Information about the two teachers was presented in the form of a rating of the candidate's performance in practice teaching. Principals were asked to examine the *Student Teacher Evaluation Form* for each teacher. One form depicted the candidate as having strong expressive but weak instrumental characteristics, while the other portrayed the candidate as having strong instrumental but weak expressive characteristics. Principals were instructed to rate the probability of success for each as a teacher in their system, and to indicate which candidate they would invite for an interview.

In order to test the two basic hypotheses of the study, a three-way analysis of variance was performed with the "Probability of Success" rating assigned by principals ($n=56$) to each teacher as the dependent variable and "System Type" (middle vs. working class), "Teacher Type" (instrumental vs. expressive), and "Level Taught" (primary vs. intermediate) as the independent variables. Principals from both types of systems clearly expressed a higher probability of success for expressive type teachers than for instrumental types ($F= 67.212$, $p<.001$). Eighty-seven and a half per cent of the responding principals indicated that they would invite the expressive type teacher for the interview.

◆ ◆ ◆

Congruence and Dissonance in the Ecology of Educational Administrators as a Basis For Discriminating Between Patterns of Leadership Behavior

GARRY FRANKLIN OLSON, Kearney State College

The study was based on the postulate derived from the theory of personality by Henry A. Murray that an individual will perform best in an environment which conforms to his personality. The problem of the study was to investigate the dimensions of interaction between educational administrators and their academic work environment which were associated with the behavior of those administrators as leaders.

Specifically, the primary purpose of this study was to ascertain the extent to which indices of the degree of congruence between personality and environmental variables could discriminate between groups of educational administrators classified according to most-effective and least-effective patterns of leadership behavior. Several levels of statistical analysis were utilized in testing five hypotheses relevant to the primary and secondary purposes of the study.

Data were collected from a random sample of 95 Illinois secondary-school principals by means of self-administered instruments mailed to the subjects and selected members of their staffs. Usable sets of data were obtained for 83 of the subjects comprising the study sample.

Ten teachers were randomly selected within each subject's school and administered the *Leader Behavior Description Questionnaire* (LBDQ) to describe quantitatively the behavior of their respective principals in the role of formal leader. Dichotomous criterion groups were established by means of scores on the LBDQ dimensions of Consideration and Initiating Structure. The most-effective leadership group was comprised of 29 subjects who scored above the mean on both LBDQ dimensions, while 24 subjects who scored below the mean on both dimensions constituted the least-effective leadership group. The data of those subjects not classified into one of the two dichotomous groups were not considered for further analysis.

The selected teachers responded to the *High School Characteristics Index* (HSCI) to describe the press of their respective school environments. The principals' responses to the *Activities Index* (AI) provided measures of their self-perceived personality needs. The discrepancy between scores on each of 30 parallel scales of the HSCI and the AI was computed to provide indices of the degree of congruence between personality needs and environmental press for each subject. These indices were indicative of the degree to which the personality of the principal matched the press of his school environment. The AI and HSCI scores, when each instrument was considered singly, provided measures for personality variables and environmental variables respectively.

Statistical analysis revealed that the degree of congruence between personality needs and environmental press significantly differentiated between the dichotomous criterion groups beyond the .01 level. The personalities of the principals considered to be the most-effective leaders were significantly more congruent with the press of their respective environment than were the personalities of the principals considered to be the least-effective leaders.

The dichotomous criterion groups were also significantly differentiated on the basis of the personality variables of AI needs scores, and on the basis of the environmental variables of HSCI press scores. Pertinent conclusions were drawn based on the statistical analyses.

Values and Value Relationships of Elementary School Principals

GALE W. ROSE, Case Western Reserve University

This paper reports part of a study on "Staff Values, Organizational Behavior, and Pupil Response in Schools", in which certain hypothesized relationships among relevant organizational variables were examined. The purpose of this paper is to present and discuss findings on the value systems of school principals and relationships of these values to certain aspects of the school environment and the school organizational behavior.

The study was based on data obtained from thirty-one elementary schools distributed over both urban and rural areas of a midwest state. In addition to the principals, 407 teachers, comprising 83 percent of the total staffs, responded. Other data provided by pupils do not enter into this paper.

Information on principals' values was secured by means of the Differential Values Inventory; principals also furnished certain demographic items and personal information. Organizational behavior was described by teachers on the School Characteristics Inventory, with indices on faculty interaction patterns and other aspects of behavior. Adequate variation among the schools was found on all major variables so that correlation analysis could be employed.

Four findings are reported for the purposes of this paper:

1. Elementary school principals whose childhood was spent in a small-town-rural environment expressed predominantly Traditional values; those who came from a metropolitan environment expressed more Emergent values. (The values construct is that of Spindler).

2. As was also found for teachers, principals who expressed Traditional values tended to locate in medium to small schools in rural areas or in lower class urban areas; those with Emergent values were more frequently found in the large middle-class urban-suburban schools. For the small, suburban elite schools no pattern was discerned for principals, although teachers tended to be Emergent.

3. In the larger schools, which were mainly in middle-class urban areas and where principals and teachers revealed predominantly Emergent values, teachers were more similar in their descriptions of behavioral interaction patterns in the school, and were more responsive to all description items.

4. Values discrepancies between principals and teachers, even substantial ones, were not related to school size, location, community-type, organizational behavior, or to principal age or sex.

The following conclusions were made from the findings:

1. There are substantial variations on organizational behavior and personal values among schools and school personnel of the same general class (i.e. the elementary school), but there are also patterns of variation related to present and past environmental conditions.

2. Principals' childhood environment appears related to present adult expressions of values, and both principals and teachers typically appear to locate in school-community settings which relate to these basic value systems.

3. Teacher variability in describing school organizational interaction patterns and in completeness of description response would appear to be in part a function of school size, and the relation of principal's values to these behaviors is ascribed to locational patterns of the principals rather than to the values themselves.

4. Differences in values between principals and teachers in the same school are not seen as relating to any predictors or effects among the personal, organizational, or environmental variables included in this study.

10:00 to 10:55 A.M. **SESSION 9.1** **DIVISION C**

TEACHER BEHAVIOR

CHAIRMAN:

Philip W. Jackson, University of Chicago

PAPERS:

The Development of a Data Bank of Diagnostic Teacher Training Information

RICHARD M. KRASNO and DWIGHT W. ALLEN, Stanford Center for Research and Development in Teaching

This paper describes the first phase of a two phase diagnostic data bank project presently being conducted at the Stanford Center for Research and Development in Teaching. The purposes of the project are as follows:

1. To collect data describing base-line teaching behavior; that is, teaching behavior exhibited prior to systematic exposure to teacher training techniques.

2. To collect data describing teaching behavior after an intensive program of training in technical skills of teaching. Comparison of these data with base-line data yields important information about the effectiveness of the teacher training techniques which were utilized (phase 1).

3. To collect data describing behavior "in the field," after the teacher trainee has taught for one year. These data provide information concerning transfer and retention of teaching skills acquired during the teacher training period (phase 2).

4. To collect data from a wide range of psychometric devices and to investigate correlations between factors derived from these tests and teaching behavior.

5. To develop a "data bank" which will function as a consortium by providing systematic access to large amounts of data on teacher training *and* by fostering communication among individuals and institutions actively interested in teacher training problems.

At the present time data collection for phase 1 of the project has been completed. All of the prospective teacher trainees (162 "interns") in the Stanford Teacher Education Program were given the following battery of tests in June 1967 during their first day at Stanford.

1. Educational Opinion Inventory
2. California F Scale
3. Kerlinger test of Teacher Attitudes

4. One-word Sentence Completion Test
5. Directed Imagination Test for Teachers

During the second and third days, the interns taught a 40-minute lesson to a class of 25-30 high school students. These lessons were video-taped in their entirety. One-third of the interns were given outlines of "pre-set" lessons in their subject area 24 hours prior to teaching, and were instructed to construct a lesson from the outline; the remainder of the interns were instructed to teach any lesson with which they felt comfortable. Ratings of the teachers on two teacher competence instruments devised at Stanford were solicited from the students immediately following each lesson. This exact procedure was repeated in mid-August 1967 immediately following the intensive summer teacher training program.

Video-tapes are presently being rated for behavior changes along dimensions of technical skills of teaching and stylistic variables. Scores on the psychometric tests are also being correlated with frequency counts of various aspects of teaching behavior. It is anticipated that this information will be valuable in selection and prediction, and ultimately will be utilized in development and modification of teacher training techniques.

◆ ◆ ◆

A Study of the Construct Validity of the Teaching Situation Reaction Test

C. KENNETH MURRAY, Northern Illinois University, JAMES K. DUNCAN, Ohio State University

This study was designed to investigate the construct validity of the Teaching Situation Reaction Test.

After examining the T.S.R.T. and consulting with the authors, certain factors were proposed as being built into the T.S.R.T. These factors were objectivity, sociability, control, confidence, reflectiveness, and empathy. Valid and reliable research instruments were selected as operational measures of these factors.

The study sought to discover if there was a relationship between the T.S.R.T., factors in the T.S.R.T., the option rankings of the T.S.R.T, and the factors previously listed.

The sample consisted of 238 pre-service education students. The data were analyzed by a 100 × 100 factor analysis program with rotation and product moment correlation.

The findings show that:

1. As an instrument, the T.S.R.T. appears to be related to objectivity, control, confidence, and empathy.

2. There appears to be a relationship among some factors in the T.S.R.T. and control, empathy, sociability, objectivity, and reflectiveness. There also appears to be some concurrent relationships between control and empathy.

3. A consistent relationship appears between the option rankings of the items of the T.S.R.T. and control. The second highest relationship is found with empathy. There are also scattered relationships with the other four factors.

It seems plausible that if you know the personality traits of models of personality that relate to performance on the T.S.R.T., it would be possible to develop personality profiles of reactions to situations which are posed in the T.S.R.T. It appears that one such profile emerges from the findings of this study.

This profile consists of a relationship between increased skill in reacting to classroom situations and higher scores on empathy and lower scores on control. A second dimension of this profile consists of an inverse relationship between skill and reaction to classroom situations and scores on control.

◆ ◆ ◆

Teacher Decision Making: The Influences of Instruction and Set on Attitudes in Reasoning

ELWOOD B. TRAYLOR, Wichita State University (Richard C. Anderson, University of Illinois)

This study examined the influence of instruction and set on decisions about the validity of categorical syllogisms whose conclusions are attitudinally relevant. Conceptually, decisions about the validity of syllogisms are viewed as the result of intellectual skills and attitudes about the conclusion. The questions being examined are: Does the improvement of an intellectual skill be carefully guided practice and/or does the induction of a set, influence the ability to make logical decisions about affective problems?

The hypotheses may be summarized as follows: (1) Carefully guided practice in solving syllogisms will improve reasoning about attitude structured syllogisms; (2) the induction of a set will improve the ability to solve attitude structured syllogisms.

140 Ss from six courses were used. Early in the semester, the MTAI was administered. Selected statements from the MTAI were used as conclusions for syllogisms constructed for the Syllogisms Test, which was administered mid-semester as a pretest to determine attitude errors. Logically valid conclusions that S disagreed with and judged to be invalid, or logically invalid conclusions that the S agreed with and judged to be valid, were

classified as attitude errors. Four weeks after the pretest the Ss were assigned at random to one of four experimental groups: programmed instruction and set; programmed instruction and no set; set and no programmed instruction; and no set, no programmed instruction. The Syllogisms Test was administered after the experimental conditions were introduced.

The data were analyzed to determine the influence of programmed instruction and set on the reduction of attitude errors. Analysis of covariance of a 2 × 2 table was used to test the hypotheses. Both hypotheses were supported.

A supplementary finding demonstrated instruction to be significant in reducing the number of non-attitude errors in deciding the validity of syllogisms. The induction of a set had no significant influence on the reduction of non-attitude errors.

◆ ◆ ◆

Relationships of Cognitive Complexity to Specific Behavioral Variables

RICHARD J. REYNOLDS, University of Georgia

The purpose of the study was to investigate cognitive complexity for its relationships to verbal interaction patterns, student attitudes, and student achievement in order to provide information for making differential predictions in the educational setting.

Student and instructor samples were selected from freshman composition sections at one university. Each instructor was given the Modified Role Rep Test (Bieri), and was observed during one semester in his teaching role. Each student was given the same test and was pretested on achievement. After fifteen weeks of instruction, each student was tested on attitude and retested on achievement.

Relationships between instructor cognitive complexity and verbal interaction patterns in their classes were examined by correlating complexity with the five major categories suggested by Flanders' interaction analysis. Student attitude and complexity were examined through correlational procedures. Student attitude and achievement were examined as functions of student-instructor cognitive style by using correlational procedures and covariance analysis respectively.

Verbal interaction patterns varied considerably in the classes of those instructors rated cognitively complex. In the classes of the cognitively simple, instructors' variation in verbal patterning was highly related to complexity level. Cognitively simple instructors spent most of their class time in lecture and direct questioning, but allowed little or no time for confusion or unsolicited student questioning. Their classes were generally more structured than those of the cognitively complex instructors in which more time was spent in student discussion and silence/confusion. A nonlinear relationship between complexity and interaction pattern suggested the possibility of a threshold effect.

The relationship of student complexity to student attitude suggested that the two variables were dependent. Much of this relationship, however, was explained by the high correlation found in one student subgroup, cognitively complex students who had cognitively complex instructors. Achievement was found to be unrelated to cognitive complexity, and not a function of the cognitive style of the students, when students are matched with their professors on cognitive complexity.

10:00 to 10:55 A.M. **SESSION 9.2** **DIVISION C**

REINFORCEMENT I

CHAIRMAN:

Charles M. Garverick, State University of New York, Buffalo

PAPERS:

Learning Principles and Procedures Applied to Remedial Reading With Normal and Educable Mentally Retarded Fourth, Fifth, and Sixth Grade Children

ADRIAN P. VAN MONDFRANS, Purdue University (Richard C. Anderson, University of Illinois, Urbana)

Arthur W. Staats has devised remedial reading procedures based upon traditional conditioning principles (Staats & Butterfield, 1965; Staats, Minke, Goodwin, & Landeen, 1967). The procedures were sufficiently simple and easy to administer so that, in the two studies cited, non-professional personnel were quickly taught to administer the program. Another important feature of the procedures was the use of extrinsic reinforcers in the form of tokens which could be exchanged for money. The money, in turn, could be used to purchase goal objects of *S*'s choice.

The present study was designed further to validate the Staats' procedure and to demonstrate their generality by extending their application to a population of pre- and early-adolescent retarded readers, including 16 experimental and 16 control *Ss* from regular 4th, 5th, and 6th grade classrooms, and 8 experimental and 8 control *Ss* of equivalent chronological age from special classes for the mentally retarded. The need for monetary reinforcement in conjunction with the token reinforcer system was also tested. The program was administered by housewives recruited from the community.

In general, the 24 experimental *Ss* participated in the study with excellent cooperation, attention, and work behaviors, thus substantiating the motivational effects of the extrinsic reinforcers. The mean number of reading responses made during the experiment was over 51,000. A mean of 550 new (unknown) words was presented with an average retention rate of 79% of these words on a short-term retention test and 54% on a long-term test.

The results showed that Ss taught by the Staats procedures, using nonprofessional administrators, scored as high on the Lorge-Thorndike Group Intelligence Test and the California ReadingTest as did the control *Ss*. Also, experimental *Ss* reinforced with tokens performed as well as experimental *Ss* reinforced with tokens and money. The experimental Ss scored significantly higher on a 100-item word recognition task, and made a smaller percentage of errors in prose reading task than did control *S*s. There were no effects of type of reinforcement upon performance in the word recognition or the prose reading task.

On those measures internal to the procedures (such as error rate, number of responses, etc.), the performance of experimental *S*s from regular classrooms receiving monetary reinforcement again did not significantly differ from the performance of the corresponding experimental *S*s receiving token reinforcement. Comparisons between the two experimental groups of educable mentally retarded *S*s did show differences in favor of those *S*s receiving monetary reinforcement on measures of response rate; but a lack of equality on pre-test scores in favor of the money-reinforced *S*s complicates interpretation of these data.

It was concluded that the procedures are generally amenable to use by these types of *S*s and to administration by non-professional personnel. The addition of monetary reinforcement to the token reinforcer system did not have significant effect on the performance of *S*s from regular classrooms, but appeared to have an effect on *S*s from classes for the mentally retarded.

◆ ◆ ◆

The Effects of Immediate and Delayed Reinforcement Treatments with Respect to Knowledge of Test Results

KATHRYN B. DANIEL, University of South Carolina

Reinforcement has been established as a major point of focus in prominent theories of learning. Evidence has, in general, supported the thesis that immediate reinforcement enhances learning to a greater degree than delayed reinforcement. However, contrary results have been found in several studies (Sassenrath and Yonge, 1967; Daniel and Witchel, 1967; Sturgis, 1966; Sturgis and Crawford, 1963; and Brackbill, Wagner, and Wilson, 1964).

This report is the second in a series of studies that investigates immediate and delayed reinforcement treatments on retention with respect to knowledge of test results. Daniel and Witchel (1967) found that scores of students who were not given knowledge of test results for one week after the test were significantly higher after two weeks on a retest than students who scored their own tests and knew immediately how they had performed on the pre-test. The findings of that study posed several questions for further investigation.

The purpose of this study was to determine if scores would remain higher for the delayed reinforcement group if the time interval for retest were changed to one week, and if the amount and type of reinforcement were changed.

The subjects in this investigation, as in the previous one, were students enrolled in educational psychology classes. Treatment was randomly assigned to two groups of 36 students each. The subject matter remained that of basic statistical concepts and was presented to both groups by the experimenter. The same 40 item examination was used for test-retest scores. The announcement of an examination over the material was made to both groups. After a four-day presentation of the material, both groups were given the pre-test.

With the experimental group, each student's test was corrected as it was turned in, and students were instructed to find correct responses for errors made. After all papers were corrected, the entire group discussed the test, and explanations for correct answers for all items were given. The experimental group had immediate knowledge of test results. One week later the same test was administered to the group.

The control group after the four-day presentation of the subject matter was given the pre-test. The students were free to leave as they turned in their papers. The next day corrected tests were returned and discussed. The control group received a 24-hour delay in reinforcement. One week later the subjects were retested.

The scores of the pre-test were tested for differences by means of analysis of variance.

The results of the present study show significant differences between immediate and delayed reinforcement groups, with greater gains achieved by the delayed reinforcement group. When special treatment with respect to the amount and type of reinforcement is given to students immediately after a test situation, the treatment seems to have no significant advantage over a 24-hour delay in knowing test results.

◆ ◆ ◆

Attitude Change Through Reward of Verbal Behavior: Replication and Extension

GLORIA ALVERNAZ MULCAHY, Drexel Institute of Technology and CHARLES H. FLATTER, University of Maryland (Doreen Steg, Drexel Institute of Technology)

Scott (1957) studied the effects of reward and punishment on the determination ot attitudes. The present study replicated Scott's experiment, and extended it to include an investigation of the relative importance of social approval and cognitive involvement in attitude development.

Subjects were junior students in education at the University of Maryland. Attitudes toward three controversial issues were assessed after exposure to a debate situation. Changes in attitude were determined utilizing pre-and-post written statements which were scored (i.e. 1 = very pro and 7 = very con). Ss who debated took the position opposite from that which they had initially written. Ss were divided into debaters and observers. Debaters were subdivided into winners, losers, and no decision. Winners and losers were predetermined although the choice was presented as being dependent upon a class vote. Winning (i.e. social approval) was deemed reinforcement for the new behavior. For replication purposes Ss were divided into three groups: 1) winners; 2) losers; and 3) observers. The extension of the study added two groups: 1) no decision debaters and 2) no decision observers. It was predicted that changes for these two groups would result from cognitive contact with opposing arguments and verbalization of the new opinions.

The results were not congruent with Scott's findings. The experience of reward or punishment or exposure to opposite opinions did not account for the differences found in attitude change.

The extension indicated that debaters receiving no reward or punishment had significantly greater attitude change in the direction opposite to their original opinions than debaters who received either reward or punishment or observers of debaters. This finding supports the idea that cognitive contact is the crucial factor in attitude change. An implication is that the factor of reward or punishment tends to interfere with the amount of attitude change. Social approval or disapproval of a new behavior tends to lead to less change in the predicted direction.

◆ ◆ ◆

Teaching Styles in Four Year Olds

NORMA FESHBACH and GERALDINE DEVOR, University of California, Los Angeles

This is one of a series of studies designed to explore differential responsiveness to positive and negative incentives and reinforcements. The focus of the present experiment is on the variations *in the use of* rather than *responsiveness* to positive and punitive reinforcement techniques.

The subject population was 204 three and four year old boys and girls enrolled in nine different Nursery and Head Start programs. The sample represents four different social class-ethnic groupings: Advantaged White, Disadvantaged White, Advantaged Negroes, and Disadvantaged Negroes. After the four year olds were taught to solve a puzzle by the experimenter, they in turn were required to teach the same task to a three year old. The primary dependent variable was the number of positive and negative reinforcements spontaneously emitted by the four year old teacher-child to his three year old "pupil".

The results were consistent with the hypothesis that advantaged white children would use positive reinforcement significantly more often when instructing their "pupils" than do their counterparts in the other three groups. Although a content analysis of the verbal statements made by the children indicates that quantitatively and qualitatively the reinforcements are similar for all four groups, the direction and affective tone of the statements distinguish the middle class white children from the other three groups. This finding is consistent with the results of previous studies but yields a different perspective—namely, at age four, middle class white children manifest a response mode which is clearly disparate from the other ethnic-social class groups. The implications of using this response mode, as a substitute measure for learning style, in determining differences between other groups, ages, and cultures, is discussed in view of the findings.

10:00 to 10:55 A.M. **SESSION 9.3** **DIVISION C**

MEMORY

CHAIRMAN:

M. David Merrill, Stanford University

PAPERS:

Effects of Grammatical and Associative Structure, Delay Interval, and Activity During Delay on Memory Span of Retarded Children

MELVYN I. SEMMEL, University of Michigan

A short-term memory deficit has been attributed to retarded children by Ellis and others. Bateman and Weather all have argued that such a deficit may account for poor language function at the automatic-sequential level of their language behavior.

This study investigated the effects of four variables upon short-term memory: grammatical structure, strength of inter-word association, length of the presentation-recall interval, and the nature of activity during the presentation-recall (delay) interval.

*S*s were first administered the digit span subtest from the Illinois Test of Psycholinguistic Abilities; subsequently they practiced naming random digit sequences from a memory drum.

During the concluding phase of the experiment, *S*s recalled one of four types of sentences two to eight words in length and differing in grammaticality and inter-word association: 1) High Association-High Syntax (HA-HS)—grammatical and meaningful sentences; 2) High Association-Low Syntax (HA-LS)—reversal of the HA-HS sentences; 3) Low Association-High Syntax (LA-HS)—replacing words in the HA-HS sentences with words of the same form class from other HA-HS sentences, retaining grammatical structure while minimizing association between pairs of words; and 4) Low Association-Low Syntax (LA-LS)—words in the HA-HS sentences were pooled and assigned to 2-8 word sentences by a table of random numbers.

Word stimuli were presented at delay intervals of 1, 6, and 20 seconds. One randomly assigned group named numbers from the memory drum during the delay interval while the other group remained silent until the signal for recall.

Each *S* received only one sentence type and either silence or number naming during the delay intervals. Delay intervals were presented across all *S*s.

Ten *S*s (ca 8 to 12 years) with IQ's from 50-80 randomly assigned to each of eight experimental conditions. All subjects were residents of Wayne County Training School, Northville, Michigan.

Significant differences were found in the main effects of delay intervals, sentence type, and type of intervening activity. An interaction between delay interval and type of intervening activity was also revealed. A significant deterioration in performance over delay intervals was found when number naming was the intervening activity but no decrease in performance when *Ss* remained silent during the delay interval.

◆ ◆ ◆

Verbal Labeling and Retention

RONALD E. JOHNSON, Purdue University

What is the effect of a verbal label upon the memory of a geometrical figure? The classical 1932 experiment of Carmichael, Hogan, and Walters, which is frequently cited as an answer to the question, suffers from many methodological inadequacies, e.g., massive figural assimilation from the use of 12 stimulus figures, the non-separation of learning from memory, and biases in ratings. In the present experiment, an elaborate series of methodological steps was followed with 259 *Ss* to insure that the two stimulus-figures were equally often normatively labeled as a "Christian cross" or as a "street intersection".

Each of 427 experimental *Ss*, drawn from the 8th through 11th grades of the white public school population of Auburn, Alabama, was individually shown a single stimulus-figure for 5 sec. and was instructed to remember the figure accurately. One third of the *Ss* were told their figure resembled a "Christian cross," another third was told the figure resembled a "street intersection," and an additional third was given no label by the *E*. After a six-months' interval, each *S* made a selection from a recognitive packet of seven forms. One of the forms was the figure originally shown to the *S*, two choices were more like a Christian cross, two more closely resembled a street intersection, and the final two choices were predictions based upon an interference theory of forgetting. After making his choice, each *S* was asked to label his figural choice and to disclose whether he had rehearsed.

Despite the long-term retention interval, 67% of the *Ss* correctly selected their stimulus-figure. *Ss* who admitted rehearsal made correct choices more frequently (72% correct) than non-rehearsers (55% correct). Analysis of the error data provided some evidence that verbal labeling at the time of learning influences the quality of retention. The bulk of the evidence, however, suggests that there is some mutual interaction between the memory of the label and the memory for the geometrical form. Overall, an interference theory, based upon the assumption that there is a nonsystematic partial forgetting of the attributes of a stimulus, predicted the majority of errors.

Retrieval Cues and Children's Memory.

MICHAEL J. A. HOWE, Tufts University (Daniel W. Marshall)

The growing awareness by psychologists of the importance of retrieval processes in human memory has prompted research into phenomena such as the "tip-of-the tongue" and "feeling-of-knowing" experiences. Children and adults may be unable to recall verbal material, not because the material is inadequately stored in memory, but because appropriate retrieval cues are not available to locate and elicit the stored items. The author (Howe, 1967) recently found that when cues were provided in the form of the verbal context in which a required word had originally appeared, recall was considerably facilitated. Differences between children and adults in the ability to recall previously learned items may be due entirely to differential consolidation, i.e., the extent to which storage of items resists interference; but there may also be important differences in the number and range of retrieval cues which are adequate to elicit a retained item. This study was conducted to discover how children of three age groups differed in the extent to which the provision of appropriate cues facilitated recall of words stored in memory. The results show significant differences between ages, both in absolute recall and in the effects of retrieval cues. Interactions are examined and discussed in the light of the theoretical problem. The findings are of direct relevance to a problem in the teaching and acquisition of foreign language vocabulary, and a related experimental study is described in brief. In this study, ability to supply the foreign-language equivalents of words was measured as a function of type of cue (auditory, visual presentation of a word, visual illustration of an item) available both during initial learning and at the time of a retention test.

◆ ◆ ◆

Delayed Information Feedback and Delayed Retention

JULIUS M. SASSENRATH and GEORGE D. YONGE, University of California, Davis

Most textbooks in psychology and educational psychology mention the superiority of immediate as opposed to delayed reinforcement of information feedback as a basic principle of learning, and point at its presumed importance in educational practice. However, recently studies by Brackbill and associates, Markowitz and Renner, and Sturgis and associates, with third graders, have generally indicated (a) no difference in rate of learning between immediate information feedback (IIF) and 10 second delayed information feedback (DIF); and (b) a small but reliable difference favoring DIF on a relearning or savings measure. Therefore, the major purpose of the present study was to determine the effect of a 10 second DIF vs. IIF on a pure measure of retention for college *Ss*. The minor purpose of the study was to determine the effect of varying selected cures during feedback.

Three hundred and eleven college *S*s were presented, via a slide projector, with 60, four alternative, multiple choice questions taken from courses in introductory psychology. One group received IIF after responding to each question. Another group received DIF 10 seconds after responding. Feedback cues were presented via the projector with either the (a) stem or no stem of the question, and (b) right and wrong alternatives or right alternatives only. In all four instances the right alternative was underlined. The design of the experiment consisted of a 2 (DIF vs. IIF) by 2 (stem vs. no stem) by 2 (right plus wrong vs. right answers) design. The eight groups were given a 60 item immediate and a five day delayed retention test.

The results from the initial exposure (before feedback) showed no reliable differences between either of the three experimental conditions. On the immediate retention test there were no reliable differences. On delayed retention *S*s receiving DIF and no stem of the question performed reliably higher than their counterpart. The results, explained in terms of verbal mediation, have implications for learning theory, teaching machines, and classroom instructions.

11:00 to 11:55 A.M. SESSION 10.5 DIVISION C

EVALUATION OF TEACHING STRATEGIES I

CHAIRMAN:

Roscoe C. Brown, New York University

PAPERS:

The Effect of Exemplary Models' Characteristics and Student Set on Resultant Student Attitudes Toward the Fine Arts

WILLIAM L. GOODWIN and IRVIN R. RUBINCAM, Bucknell University

Purpose. To investigate the effect of exemplary models' characteristics and teacher-induced set upon the attitudes of school-age subjects toward the fine arts.

Procedures. Four related experiments in separate subjects were conducted under realistic classroom conditions; and Ss, public and parochial school pupils enrolled in a summer instructional program, were randomly assigned to treatments. Experiments in music and dance utilized live performing models in the classroom; while experiments in creative writing and art utilized samples of art products and fictitious models. Different independent variables were incorporated in the design of the several experiments. In the two experiments involving music and dance, similar designs were employed. The music experiment involved a 2 × 2 × 2 complete factorial design (2 levels of model competence, high and low; 2 levels of teacher-induced set, positive-prolonged and routine-short; and 2 grade levels, intermediate and junior high). The dance experimental design was identical except for adding a third grade level (primary). The third investigation, in creative writing, was a complete 3 × 2 factorial (3 treatments, no information about the model, model thought to be older, model thought to be the same age; and 2 grade levels, primary and intermediate). The fourth experiment, in art, was a complete factorial 5 × 3 × 2 design (5 treatments, no information about the model, model the same sex and same age as *S*, model same sex as *S* and adult, model opposite sex as *S* and adult; 3 grade levels, primary, intermediate, and junior high; and sex of *S*). The dependent variable used was attitude toward the subject field, as measured on locally-developed instruments.

Results and Conclusions. Analyses of variance revealed model competence to be significant in the music experiment and grade level significant in the dance experiment; teacher-induced set was non-significant. No main effect was significant in the art or creative writing experiments. Results are discussed in regard to differences between (as well as within) each experiment.

Evaluative Teaching Strategies for Environmental Contamination Topics in the Social Studies

MILTON MEUX, KEITH EVANS, GEORGE ENDO, and
MICHAEL HOGBEN, University of Utah

Smith, Meux,. Coombs, and Nuthall have developed a system for analyzing teaching strategies in the classroom discourse in the main high school subjects. The strategies are composed of moves, small units of information relevant to achieving the objective or central point of the strategy.

The purpose of this study is to investigate the effects of four evaluative teaching strategies—strategies in which the objective or point is to rate some action, practice, or event. The strategies vary in degree and kind of support for the rating of an environmental contamination topic, either fluoridation or use of pesticides, as inadvisable. The first strategy contains only Describing moves. The second strategy contains this core of Describing moves plus Citing-Authority moves. The third strategy contains the core plus Analogy moves. The fourth strategy contains the core plus Criterion moves. Comparable tests were constructed for each topic, with both multiple-choice and free-response items.

The four strategies were assigned randomly within each of several classes of eleventh-grade Social Studies students (N = 303), each student receiving comparable strategies for each of the two topics. Topic order was counterbalanced, along with the order of the corresponding tests for each topic. An eight-scale semantic differential for each topic was administered before and after each strategy.

On multiple-choice items, the Criterion group performed somewhat better, especially selecting reasons against fluoridation or pesticides. On free-response items, all groups indicated more concern with harmfulness or effectiveness than all other response categories, with harmfulness predominating in pesticide items (but some Ss defend pesticides as necessary) and effectiveness in fluoridation items (although the ratio of harmfulness to effectiveness responses increases with degree of support). The Criterion group performed somewhat better overall. On the semantic differential, the only significant changes were on the successful-unsuccessful and safe-harmful scales for the Criterion group.

◆ ◆ ◆

The Teaching of Inquiry Skills to Fifth Grade Children

JEROME S. ALLENDER, HEDY ZUSSMAN, DONALD R.
DUTTER, and EDWARD S. JUROWSKI, Miami University

The main question of this study was to determine whether fifth grade children can be taught to increase their inquiry activity. Previous work by Suchman has shown that the number of questions elementary school children ask can be increased, but the conditions under which they are increased are equivocal. The inquiry materials used (entitled. *I Am the Mayor*), allowed the children to play the role of a mayor of a simulated small city. Measures of problem sensitivity, problem formulation, search behavior, and inquiry time were made on 54 children, three groups of 18, for four separated trials. The children in the control group attended their school's regular learning center on days between trials. The two experimental groups attended a learning center designed to teachinquiry skills under two degrees of structure: for the first group the program was teacher-directed and for the second group the program was self-directed. It was hypothesized that mean increases in problems looked into, questions asked, information requested, and inquiry time would be (1) least for the control group, (2) of middle value for the group that attended the learning center whose program was structured by a teacher, and (3) greatest for the group that attended a learning center whose program was structured by the student. All of the mean scores for the four measures, with one exception, were in the predicted direction. The differences between the experimental and control groups were significant at the .10 level. The major conclusion was that inquiry activity is positively effected by unstructured teaching.

11:00 to 11:55 A.M. SESSION 10.8 DIVISION E

COUNSELING; EVALUATION AND MICRO-PROCEDURES

CHAIRMAN:

Isadore Goldberg, Computer Applications Incorporated

PAPERS:

Adjunct Roles of Data Processing Devices in Computer Assisted Counseling

TOMMY L. ROBERTS, Oklahoma State University and FRANZ J. FREDERICK*, Purdue University

This study considers the problems of (1) quantity and quality of data available to public school counselors; (2) limited student access time to counselor; (3) lack of adequate conceptualization concerning student decisions; (4) lack of adequate conceptualization of counselor role in the student decision making process; (5) lack of qualified counseling personnel to help in student decision making processes and the constraints these problems impose on human counseling systems.

A cluster of ten major student decisions was identified and the "decisions" dichotomized as present imperative and future imperative decisions.

Methods of facilitating student decisions have been proposed which range from on-line interactive computer sequences to non-directive counseling. Previously proposed man-machine systems have however implemented only one or two of the present imperative decision problems.

In an attempt to develop a comprehensive man-machine counseling system, a taxonomy was postulated in the form of a two way classification table with decision control and operating system for information accessing as the major dimensions. The taxonomy maps a domain of facilitation of student decisions, and allows consideration of the various roles of high speed data processing devices both on-line and off-line ranging from student interactive to student-counselor interactive. The domain limits are predicated upon the data processing devices functioning in *adjunct* capacities.

A pilot man-machine counseling system which is student-counselor interactive in nature is proposed for the facilitation of the cluster of ten major student decisions. The pilot system is progressively developed in two stages—(1) off-line and (2) on-line.

The proposed system thus retains the characteristics of present counseling systems while capitalizing upon the various advantages afforded by high speed data processing devices utilized in adjunct capacities.

Project Self Discovery: Its effects on bright underachievers at 9 Florida high schools.

WILLIAM W. PURKEY, University of Florida (T. Antoinette Ryan, Oregon State University)

The purpose of the study here reported was to produce and test an independent study project designed to enhance the personal and social adjustment, and to improve academic achievement of bright but underachieving high school students.

For 8 months, 44 bright but underachieving 10th and 11th grade students from 9 Florida high schools used a specially prepared "Project Self Discovery" manual designed to enhance student self concepts and improve academic achievement through independent readings and written exercises submitted to the researcher for comment and return. The California Psychological Inventory (CPI) and a Self Ranking Inventory (SRI) were administered before and after treatment, and results compared with those of a control group of 40 taken from the same pool.

While GPA's of both E and C groups increased, few changes occurred in SRI or CPI scores. Where changes occurred, CPI and SRI increases were as frequent in the C group as in the E group, but on different CPI and SRI characteristics. Numerous SRI and CPI changes were found among subgroups. Rural S's were affected differently from urban S's , boys differently from girls, and 10th graders differently from 11th graders. All signifiant changes were in a positive and favorable direction.

◆ ◆ ◆

A Subjective Measure of Objective Success in Group Counseling—A Longitudinal Analysis

MILTON D. JACOBSON, University of Virginia and VICTOR G. HADDOX, M. D., Queen's Hospital, Honolulu, Hawaii

The purpose of this study was to determine if the emotions of counselor and counselees are differentially effected by their interaction in group sessions occurring over a three month period. If differences occurred, these would be related to the success of roles played during the sessions.

A group of six volunteer subjects were given "Mood Scale Cards" developed by Dean J. Clyde to determine the subjects' emotional state prior to and during hypnosis. In this testing device, a subject arranges a set of 48 cards according to his *subjective* feelings. A rating of six different emotions is obtained. Group trance induction was then initiated. Trance depths sufficient to produce positive and negative hallucinations were achieved. The subjects were asked to complete various tasks while in this state, all of which were emotionally nonprovocative and nonthreatening. The subjects were again presented with the Clyde Mood Scale Cards. Then the trance was terminated. Thus an evaluation was made of the subjective emotions of the subjects in a normal (non-trance) and in a deep hypnotic state. The experimenter, who was not in a trance condition, sorted Mood Scale Cards whenever the group did.

The data obtained were analyzed by variance and covariance techniques. Results showed that the emotions of the subjects did not change but that the emotions of the counselor changed significantly during the sessions and over the three month period. Initially, the counselor's emotions were atypical of him. Large changes occurred in beginning sessions, but changes became minimal after several sessions.

The counselor's emotional changes were related to his success with the experiment as determined by several *objective* analyses, and by the reactions of the subjects (counselees).

In this investigation the success of the clinician was shown in two ways—by the relaxation of his own emotions from a tension to a normal state, and by the reflection of this success in the emotions of the group with which he is working. The Clyde Mood Scale Cards provided a means of evaluating this success.

◆ ◆ ◆

Comparison of Three Experimental Modes of Counseling

GARTH SORENSON and RICHARD HAWKINS, University of California, Los Angeles (T. Antoinette Ryan, Oregon State University)

In a quasi-laboratory setting, three experimental modes of counseling were compared on the basis of three interviews with regard to their effects on behavior, moods, and feelings about counseling.

In mode 1, "inquiry", counselors guided subjects through a series of steps designed to help them make plans for coping with their problems. In mode 2, "directive", counselors made brief case histories and suggested specific actions to counselees. In mode 3, "affect", counselors encouraged exploration of feelings.

Each mode was defined by an interview schedule which specified several verbal categories of counselor behavior, and rules about when each category should be used. Three experimental counselors were trained to use all three modes. Following the interviews, two trained raters, working from tape-recordings of the interviews, confirmed that counselors had followed the modes prescribed.

The subjects, 36 student teachers, all women, who reported more than average stress in student teaching, were classified according to academic ability and relative stress. They were randomly assigned both as to counselor and counseling mode. Reports of their reactions to counseling were collected during and following the three interviews.

The study shows that

1. low stress counselees were more likely to drop out of counseling than high stress counselees;

2. subjects in mode 2, "directive", were more satisfied with counseling than were subjects in other modes;

3. more subjects in mode 1, "inquiry", tried out plans made during interviews than did subjects in the other modes;

4. subjects with low grades tried more alternative plans of action than did subjects with high grades;

5. of the subjects assigned to mode 1, "inquiry", those with high grades were more satisfied with counseling than were those with low grades.

Conclusion is that different modes of counseling are needed for different kinds of students and different counseling goals.

11:00 to 11:55 A.M. **SESSION 10.9** **DIVISION C**

VERBAL BEHAVIOR III

CHAIRMAN:

Robert E. Davidson, University of Wisconsin

PAPERS:

Word Recognition and Discrimination Errors by Young Children

JAMES W. HALL, Northwestern University (Richard C. Anderson, University of Illinois)

It has been shown that when a familiar word (e.g., GALLOP) is perceived, strong associates of that (e.g., HORSE) are likely to occur as implicit responses, and that such occurrence frequently leads to a "false recognition" of the implicit response. That is, the implicit associative response or IAR (HORSE) is falsely identified as having been the perceived word (GALLOP). Such false recognitions are more common among five-year-olds than among eight-year-olds, leading to the hypothesis that the older *S*s are better able to discriminate between words presented to them and words elicited as implicit responses by the presented words. The present experiment tested a prediction based on this hypothesis.

Each of 27 first- and 27 fourth-grade children were presented aurally 20 familiar words termed Stimulus Words. Each *S* was asked to respond aloud to each Stimulus Word with the first word (Response Word) that came to his mind. Next, the 20 Stimulus Words, the 20 Response Words, and 20 "New Words" were shuffled together and presented in random order, with the *S* instructed to specify in which of the three categories each word belonged.

The discrimination hypothesis suggested to account for the decrease with age in false recognition rate was supported. In instances where a word was correctly recognized as having occurred previously, the younger *S*s were more likely to fail to discriminate between its appearance as a Stimulus Word vs. as a Response Word. This effect was particularly great in the case of words that *S*s had given as responses, the case presumed to be analogous to the occurrence of IAR-produced false recognitions.

Also of interest was the fact that Response Words were misclassified as New Words three times more frequently than were Stimulus Words. That is, the experimental procedure resulted in a markedly lower level of learning for the Response Words despite the fact that those words and not the Stimulus words had been pronounced aloud by the *S*s.

Vowel-Pairs: Domination, Cooperation, Creation, Separation

LOU E. BURMEISTER, University of Wisconsin

This study was designed to investigate the phonemic behavior of single grapheme vowel-pairs in order to find if it is possible to formulate valid phonic generalizations for vowel-pairs in commonly found English words. Such generalizations at present are not available.

The findings reported in *Phoneme-Grapheme Correspondence as Cues to Spelling Improvement* were reversed to grapheme-phoneme correspondences in order to describe the sounds of all single grapheme vowel-pairs in the 17,310 words which Hanna, *et al.*, examined. The frequency and percent of occurrence of all phonemes for each grapheme are reported.

To analyze the data, the author grouped all vowel-pairs which occurred at least 50 times in the sample material according to each grapheme's most common phoneme. General groupings are: digraphs, diphthongs, and others. Within this framework, two types of findings are reported: percent of occurrence of the three most common phonemes, or the number of phonemes needed to reach the 90% utility level (to cover 90% of the occurrences of the grapheme), whichever occurs first, for each grapheme: frequency of occurrence of same.

1. Digraphs: The first vowel usually dominates in the following graphemes: ai, ay, ea, es, ca, ow. Second sounds should be taught for "ea" and "ow," on the basis of both frequency and percent. No other vowel-pairs are digraphs.

2. Diphthongs: The two vowels usually cooperate (blend) in the following graphemes: au, aw, oi, oy, oo. "Oo" has two common sounds. "Ou" is the only other diphthong, but its most frequently found sound is a schwa.

3. A new sound may be created by ie (ā), ou (e), ey (i), ew (ū). These are the respective graphemes' most common sounds.

4. Except for the most common sound of "ei," "ei" and "ie" have the same sounds, in the same order of occurrence: ē, i, ī.

This study furnishes proof that some vowel-pair phonic generalizations which are commonly appearing in the literature are not only fallacious but could be replaced by valid generalizations. Further, this study provides such generalizations as well as proof of their validity.

◆ ◆ ◆

Effects of Discrimination and Reproduction Training on Ability to Discriminate Letter-like Forms.

JOANNA P. WILLIAMS, University of Pennsylvania

Maccoby and Bee (1965) hypothesized that *S* must take account of more of the attributes of a form in order to reproduce it than to discriminate it from other forms. This suggests that reproduction training could lead to superior discrimination by forcing *S* to attend to more criterial attributes. However, discrimination training involving minimally different stimuli as negative instances should also do this, so that the extent of the superiority of reproduction training should depend on the similarity of the standard and the comparison stimuli in the discrimination task to which it is compared.

32 kindergarten pupils were given Gibson et al. (1962) letter-like forms as standard stimuli and four transformations, as well as twelve additional forms.
Discrimination training included delayed matching-to-sample

Group 1: Standards vs. right-left reversal and 180° rotation transformations;

Group 2: Standards vs. up-down reversal and 90° rotation transformations;

Group 3: Standards vs. highly differentiated forms.
Reproduction training for

Group 4 consisted of tracing and copying standards (*S*s yoked on training time to *S*s from the other groups).

Testing was of three discriminatiion tests of increasing complexity, involving the standards of all types of comparison stimuli, administered after training, and 24 hours later.

Overall performance of groups 1 and 2 did not differ. These groups were superior to the other two groups, which did not differ. There were significantly more errors on the R-L reversal than on the other three transformations. However, the difficulty of specific transformations other than the right-left reversal was significantly greater for one-third of the standards.

Results are interpreseted as supporting in general the notion that the relative effectiveness of the training methods depends on the number of distinctive features attended to and abstracted during training. Gibson's hypothesis that the transformation types are more important as "predictors of identifiability" than are the characteristics of the standard itself, is not borne out.

SESSION 10.10

11:00 to 11:55 A.M. **DIVISION C**

DIMENSIONS OF INDIVIDUAL DIFFERENCE

CHAIRMAN:

Wai-Ching Ho, Educational Research Council, Cleveland

PAPERS:

A Factor Analytic Study of Listening

DWIGHT L. FRESHLEY and HARRY E. ANDERSON, Jr.,
University of Georgia

The present study was undertaken on the assumption that listening could probably be explained on the basis of logical reasoning, recall, vocabulary, and comprehension. A sample of 357 fifth graders from a suburban school district was administered the Opposites, Similarities, Analogies, Delayed Recall, and Verbal Comprehension subtests of the 1963 Edition, Level 2, of the California Test of Mental Maturity (CTMM Short Form); the Immediate Recall subtest from the CTMM (Long Form); and the Vocabulary subtest from the Cooperative School and College Ability Tests (SCAT), Form 5B.

The items from the above tests were intercorrelated by the phi coefficient method, factor analyzed by principal components with unity in the diagonal elements, and rotated by the varimax routine. Elimination of low-loading items provided for the definition of four relatively clear factors as follows: Verbal, with three CTMM Comprehension and four SCAT Vocabulary items; Opposites, with six CTMM Opposite items; Recall, with seven CTMM Immediate Recall items; and Similarities, with five CTMM Similarities items. A re-analysis of these 25 items using tetrachoric correlations resulted in the same four factors which were used, also, as "marker" variables in further analyses.

The students were administered the Listening subtest from Form 4B of the Cooperative Sequential Tests of Educational Progress (STEP). The 80 items of the Listening test, together with the above 25 items, were correlated by the tetrachoric method and factor analyzed as before. The four marker variables were clearly defined with numbers of Listening items loading (i.e., greater than .25) as follows: 24 on Verbal; 14 on Similarities; three on Recall; and none on Opposites. In addition, 26 Listening items loaded greater than .25 on a separate factor, while 20 Listening items had loadings less than .25 on all five factors. The discussion of results centers on item content.

Dimensions of Problem Behavior in the Elementary School

BEEMAN N. PHILLIPS, The University of Texas at Austin

School problem behavior is generally considered significant both in terms of adaptation to school and in terms of personality development; and since Wickman, overt, discrete aspects of pupil behavior have been the subject of continuing research. About 600 fourth graders were studied as part of a USOE-supported project; and data from standardized tests, self reports, teacher nominations, peers, and cumulative records were gathered at the beginning and end of both fourth and fifth grade. Utilizing a set of 72 problem behaviors, nominations from teachers on each of these four occasions were combined and factor analyzed with Kaiser's image analysis technique. Five dimensions of problem behavior were obtained, including aggression with independent strivings (AI), active withdrawal (AW), emotional disturbance with depression (ED), self enhancement through derogation of others (SE), and diffuse hyperactivity (DH). And, correlations between these aspects of problem behavior and measures of school motivation (SM), teacher grades (GPA), peer status (PA and PR, only by same-sex peers), school achievement (MAT), and academic aptitude (CTMM) were computed for middle class white, lower class white, Negro, and Mexican-American subgroups. Relationships were highest for Negro and middle class white subgroups, and lowest for the Mexican-American subgroup, with the lower class white subgroup falling in between. Plausible interpretations of these findings, including a congruity-discongruity of behavioral and educational *aspirations* hypothesis, are advanced; and a number of specific findings, including intriguing results for SM, PA, and PR, and the research advantages of the teacher as a non-intrusive observer, are discussed.

◆ ◆ ◆

The Effect of Special Instruction upon Test Performance of High School Students in Tennessee

S. O. ROBERTS, Fisk University, and DON B. OPPENHEIM, Educational Testing Service

Previous studies of the effect of coaching on test scores have been essentially negative. The present study was undertaken to determine whether students who had received less adequate instruction and who were less able in scholastic aptitude test performance than the subjects in previous studies, might benefit more from special instruction.

Specially prepared linear programed materials, designed to foster an analytic approach to the tasks required on college admission tests and to be used in conjunction with classroom instruction, were developed in the verbal and the mathematical areas. Instruction in either of the two areas was given to eleventh grade students in 18 predominantly Negro high schools. Alternate forms of the PSAT, administered as pre- and posttests, were used to evaluate the effectiveness of the special instruction.

While there were statistically significant differences in gain scores between some of the experimental groups and their controls, the outcome of this study, like the outcome of earlier studies, was generally negative. In fact, the gains in the experimental groups were more than offset by reductions in posttest scores in the control groups. The magnitude of the gains observed in the experimental groups was so small, that it does not seem reasonable to expect that similar short-term instruction given on a wide scale would be of significant benefit to disadvantaged students.

◆ ◆ ◆

Birth Order Categories as Predictors of Select Personality Characteristics

THOMAS S. ALLMAN and WILLIAM F. WHITE, University of Georgia

Past research dealing with the types and characteristics of relationships between birth order and personality factors has been relatively brief and inconclusive. This study was designed to extend and help clarify the present body of knowledge, and more specifically to examine the relationships between selected birth order rank categories and personality behavior as measured by the *Sixteen Personality Factor Questionnaire, Form A* (16PF).

The 16 PF and a questionnaire developed by the authors designed to determine birth order rank were administered to the sample of 87 female students at the University of Georgia. A factor analysis was obtained from the responses on the 16PF. From a rotation of the factor matrix emerged six second order factors. The mean scores of each of the selected birth order groups on each of the six second order factors were then compared by means of Duncan's New Multiple Range technique with correction for unequal N, and significant differences were noted.

Analysis provided significance between several of the birth order groups on two of the six second order factors. Significant differences between First Born Only and Second Born and Youngest subjects as well as between First Born Only and Third Born subjects were found on Factor V, which has been subsequently described as a sensitive-effeminate vs. tough-realistic continuum. On this continuum, First Born Only subjects fell nearer the positive or sensitive-effeminate polar meaning; Second Born and Youngest and Third Born subjects were closer to the negative or tough-realistic pole.

Significance was also found on factor VI between First Born Only and Second Born with Younger Siblings subjects. The authors have labelled this continuum conservative, simple, accepting vs. jealous, sophisticated, radical. First Born Only subjects were found to be nearer the negative or conservative, simple, and accepting polarity of the continuum, and Second Born with Younger Siblings nearer the positive or jealous, sophisticated, radicalism dimension.

11:00 to 11:55 A.M. **SESSION 10.11** **DIVISION B**

CONTRIBUTED PAPERS

CHAIRMAN:

John M. Mickelson, Temple University

PAPERS:

Review of Research related to Training for Research in Education

CORAHANN OKORODUDU, The Ohio State University

The purpose of this paper is to summarize major trends found in a comprehensive review of studies pertaining to the training of research personnel in education and related fields (Bargar, Okorodudu, et al., 1967). The primary objective of the review was to provide a propositional summarization which would have immediate policy and administrative value for the institutionalization or evaluation of research training programs. Thus, the review was generated by the need to base attempts to develop potential research (R) and development (D) talent upon valid assumptions about the relationship of specific aspects of academic settings and programs to post-academic research and development output. As a second objective, the review could also provide the basis for the identification of questions concerning what we need to know about training for R and D careers in education, as well as the projection of feasible research strategies in relation to these questions.

The review presents rather impressive evidence of the inadequate production of R and D manpower by schools of education. Major criteria in this regard are (1) the proportion of doctoral recipients who become primarily engaged in R and D activities after graduation and (2) their research productivity.

The principal manipulative antecedent factors identified are aspects of the institutional setting and training provisions of schools of education. With regard to institutional factors, it is shown that probably the most critical antecedents in the production of researchers are (1) institutional selectivity, which includes the recruitment and selection of talented students and also faculty recruitment; (2) institutional climate for research, reflected mainly by strong research emphasis and high quality of research produced. With regard to training provisions, findings pertaining to research outcomes of doctoral study in education are examined. These findings point to the relative significance of various aspects of doctoral training programs, such as degree differences in professional research orientation, the existence of specialized training for research and non-research course offerings, apprenticeships, continuity of study, and the dissertation. Attention is also devoted to the degree of predisposition toward inquiry at the undergraduate and master's levels.

Finally, findings are presented concerning several background factors classifiable as inputs which relate to research productivity via developmental channels. These factors include personality, socioeconomic status, religion, and geographic origin. Although background factors are not directly subject to manipulation in the production of researchers or developers, these factors should be considered both in the recruitment of potential research-development talent and in the provisions of a climate of creative nurturance for their fullest development.

◆ ◆ ◆

Future Directions for the Training of Researchers in Education

ROBERT R. BARGAR, The Ohio State University

The purpose of this paper is to discuss major implications drawn from an intensive review of research related to the training of educational researchers (Bargar, Okorodudu, et al., 1967). One serious problem outlined by the studies reviewed concerns the present critical shortage of research and development personnel. Under the impetus of Title IV of ESEA, schools of education are now moving more strongly in the direction of training such specialists. It seems clear from the available evidence that these efforts, though important, will fall far short of satisfying the current need. Furthermore, it is doubtful that such programs by themselves will ever resolve the supply-demand gap. The evidence reviewed suggests that the problem being faced is much deeper than a failure to produce specialists, and that alternative training routes pursued along with the production of full-time specialists will be vital to the solution of the manpower shortage.

It is the author's conclusion that the failure of universities to produce R & D specialists is only a symptom of a serious failure of most educational practitioners, and many researchers, to understand the critical relationship between inquiry and practice. As long as this situation exists, educational development will continue to be a frustrating and stumbling process. The long-range solution to this problem will require the providing of appropriately focused R & D training for educational practitioners at various academic and in-service levels. In addition to providing relevant training for practitioners, such programs will also constitute a natural means of recruiting talented educational personnel into R & D specialist roles and training programs.

The major purpose of this paper is to review evidence that supports this analysis, and to suggest types of training programs which appear possible at various educational levels. Relevant research strategies will also be discussed.

◆ ◆ ◆

Factor Analysis Applied to Binary Coded Questionnaire Data in a Vocational Education on-the-Job Survey

HERBERT GARBER, University of Connecticut

A study of new office workers employed throughout the state of Connecticut attempted to gather data relatable to the high school training the subjects had recently completed. A major aim was to "—to identify—the knowledges, skills,—and understandings (KSU) they (new workers) need in order to perform (their)—work."

A very comprehensive, 170 item questionnaire was responded to by the over 500 new employees. This huge data matrix was comprised of yes-no (zero-one) responses, and needed to be reduced to manageable dimensions. One reasonable assumption could be made from a prior knowledge of general business office practice: that the array of jobs within any office ought to show some occupational clustering which would be revealed by the patterns the interviewees created in their responding to the KSU items in the overall list.

Factor analysis with varimax rotation was chosen as the appropriate analytic technique. Since total score could not be regarded as a criterion, and no external criterion was available, a suitable intercorrelation coefficient needed to be chosen. The phi coefficient is an expression in compact form of the concordance between two variables. It requires no assumptions of population normality nor even continuity.

The analysis was performed on a correlation matrix comprised of phi coefficients derived from 94 frequently-responded to items covering such areas as typing, filing, communications, etc. The resulting factor matrix yielded 27 rather clear factors including an unexpected "honesty in responding" factor.

Psychological interpretation was rather easily made of all 27 factors, thus describing in understandable terms the work demands perceived by a very large population of workers. At the same time, business training curriculum offerings can now be re-evaluated on the basis of solid evidence about the tasks trainees have reported from direct experience.

◆ ◆ ◆

Education Programs in Prototype Prisons For Youths and Young Adults

CALVIN B. MICHAEL, Eastern Michigan University

This paper describes a study of the assumptions and characteristics of educational and vocational training programs in four prototype prisons for youths and young adults. Inmate populations were also studied and described.

The investigation resembled an anthropologic field study. Documents and paraphernalia were analyzed and checked against practices and perceptions. Staff members and inmates were interviewed and personnel records cross-checked. Studying the whole "society" while living in the institutions, the investigator also observed particular programs in operation. Although the investigation emphasized education programs, the entire institution was studied and interrelationships of persons and programs identified and analyzed.

The four institutions housed young men who typically had histories of school failure, educational retardation, and an absence of any saleable skills. Nonetheless, general and specific abilities and aptitudes appeared to be similar to the general population distribution. The inmate groups differed from one another in a somewhat hierarchical way. Successively, the four prisons contained convicts who, on the average, were slightly older, had committed more serious offenses, had more extensive criminal histories, and had had more difficulty in institutional adjustment.

Programs of education and vocational training were largely "transplants" from public schools or industrial settings. Well-publicized "innovations" tended to be adopted uncritically and without consideration of the previous school experiences and consequent attitudes of inmates. Materials were typically outmoded, and methods of instruction uninspired. Physical facilities were adequate, but their use was unimaginative. Many staff members were cast-offs from public schools, and the general level of professional wisdom and competence was inordinately low.

Most practices seemed based on naïve assumptions about the role of education and vocational training in prisoner rehabilitation, e.g. that providing vocational skills or literacy training *per se* transforms criminals. Few aspects of the programs were well articulated and seldom constituted an integral dimension of a total institutional program of treatment. It might be said, thus, that correctional education in these prisons is as deficient in its sociology and psychology as in its economics.

12:00 M to 12:55 P.M. **SESSION 11.1** **DIVISION C**

LEARNING FROM PROSE II

CHAIRMAN:

Aaron S. Carton, State University of New York, Stony Brook

PAPERS:

Some Unpredicted Effects of Difference Questions upon Learning from Continuous Discourse

LAWRENCE T. FRASE, Bell Telephone Laboratories, Murray Hill, New Jersey and JEAN PEZZOLI*, University of Massachusetts

Some questions seem to be more useful aids for study, but just what makes a "good" question is open to exploration. In the present studies it was predicted that general pre-question (which ask for several facts) would require that *S*s process more information from a 36 word passage than would less general questions (asking for a specific fact, or the comparison of two facts).

In the first study, designed to see if *S*s agreed that the questions required more or less information, 60 college *S*s were asked to underline all the words in a 36 word passage (giving names, dates of birth, and occupations of four men) which they thought were necessary to answer a specific, comparative, or general question. There were 20 *S*s who saw one of the three types of questions. In spite of pre-training, *S*s given the general questions ignored words from the stimulus portion of the S-R pairs presented in the paragraph. A Kruskal-Wallis analysis of variance indicated that the general question group had a significantly higher number of stimulus extrusions than either the specific or comparative question groups, even though the number of words seen as relevant increased from specific to general questions. Neglecting critical stimuli was considered to be an information rejection strategy which accounted for the results of the second study.

In the second study 84 *S* were instructed to use the questions as aids in learning the 36 word passage. There were 28 *S*s under each of the question categories. Subjects were allowed 20 seconds to study the passage above which they saw the question. An analysis of variance showed that retention (measured by a multiple-choice test) decreased from specific to general groups whether the criterion was retention of the whole passage or retention of a specific piece of information relevant to the most precise question. The results thus contradicted the original hypothesis.

Learning from connected discourse seems to involve both a response learning and associative phase. Questions which are to be used as instructional aids must be phrased in such a way that *S*s are directed to rehearse the stimuli, the responses, and also the associations between the two.

The Roles of Pictures and Readability in Comprehension of the Main Idea of a Paragraph

KARL KOENKE, University of Wisconsin

The study was designed to answer two general questions: (1) Does the addition of content relevant pictures to expository paragraphs enhance main ideas evoked from third and sixth graders? (2) Does simplification of expository paragraphs according to readability formulae enhance main ideas evoked from third and sixth graders? The effects of direction to attend to pictures were also considered.

Three main ideas were conceived and each was expressed verbally, in a paragraph, and pictorially. Each paragraph was written at three levels according to readability formulae—basic, third grade, and sixth grade. Main ideas were elicited from the 192 subjects after they had read three paragraphs alone, with pictures, with pictures and directions to view the pictures, or with pictures and directions concerning the picture-text relationship. At least two of the three judges, who used a specially developed 7-point scale, gave identical ratings to 96.7% of the responses. A 4 × 2 × 2 × 2—Reading Conditions, Readability Levels, Grade, and Sex—completely crossed factorial design was used.

Analyses yielded the following results: Reading Conditions and Sex main effects were not significant. Significant Readability and Grade effects and interactions revealed that only sixth graders' response ratings were significantly higher if simplified material was read, and only boys' response ratings were significantly higher in sixth grade than they were in third grade.

The conclusions were: The addition of a picture to an expository paragraph, with or without direction to use the picture as a comprehension aid, does not enhance children's main idea responses. Varying the readability level from that equal to the reader's attained grade to an extremely simple level, enhances the main ideas elicited from sixth graders but not third graders. Although some researchers have found that boys make better use of pictures as aids to reading comprehension, the present data do not support this conclusion.

◆ ◆ ◆

Effects of Intonation upon Listening Comprehension

ANDREA RICE ROZRAN, University of Chicago

Some authors have asserted that the intonation features of language—that is, pitch, stress, and juncture—carry information which is essential for complete comprehension of oral language. They argue that comprehension cannot occur unless the phrase structure of language is perceived and that pitch, stress, and juncture are the principal devices for signalling phrase structure.

This study investigated whether the intonation features of language carry information which is essential for complete comprehension of oral language.

The hypothesis was tested by selecting two passages, each approximately 250 words in length, at the Dale-Chall 4.5 and 6.5 readability levels. Each passage was tape-recorded in both list intonation and full expression intonation.

Two randomly-selected groups of fourth grade children heard recordings of the two passages in such a way that each group heard each passage only once and each intonation pattern only once. They answered comprehension questions over each passage. There was one testing session for each group. A 2 × 2 factorial design was used, with a control for order effects.

The first hypothesis stated that passages read with list intonation were more difficult to comprehend than those read with full expression intonation. An F of .77 was found, which was not significant at .05. The second hypothesis stated that there was an interaction between the intonation in which a passage was read and the readability of that passage. An F of 6.99 was found, which was significant at the .05 level.

Using the described materials and procedures, list intonation appeared to aid the comprehension of difficult passages while it hindered the comprehension of easy passages.

The theory that intonation is necessary for complete comprehension of verbal materials was placed in doubt, but it cannot be contradicted until the study is replicated in attempts to develop a specific theory to explain the observed results.

◆ ◆ ◆

Using Multiple Choice Questions to Measure the Effect on Comprehension of Material Written with Certain Oral Language Patterns

SUSAN M. TATHAM, University of Wisconsin

Two recent studies (Ruddell, 1963; Tatham, unpublished, 1967) have shown that children comprehend passages that are written with patterns that appear frequently in their oral language significantly better than passages written with patterns that appear infrequently in their oral language. In both studies, however, the close procedure was used to measure comprehension. Because it was felt that deletion of words in unfamiliar sentence patterns complicates the reading task more than deletion in familar patterns, the present study was conducted to answer the following question: Is comprehension of a passage written with frequent language patterns significantly greater than comprehension of a passage written with infrequent language patterns when comprehension is ascertained by multiple choice questions?

Fifty-three fourth graders were randomly assigned to two treatment groups stratified on the basis of sex and IQ. Group IA read a 254-word passage that used language patterns identified by prior research as occurring frequently in fourth graders' oral language (Strickland, 1962); Group IB read the same passage but with language patterns identified by prior research as occurring infrequently in fourth graders' oral language (Strickland, 1962). Vocabulary level and sentence length were controlled by the Dale-Chall Readability Formula. Comprehension was determined by multiple choice questions.

In contrast to previous research, results of the unequal cells analysis of variance indicated that comprehension scores on the passage that used frequent language patterns were not significantly greater than on the passage that used infrequent language patterns.

If the findings of this study can be confirmed with more rigorously tested materials and with diverse populations, two basic conclusions might be reached: (1) the close procedure does not measure the same aspects of reading comprehension as multiple choice questions over materials written with select language patterns; and (2) the use in written material of frequent or infrequent oral language patterns does not affect the accuracy with which children read.

12:00 M to 12:55 P.M. SESSION 11.3 DIVISION A

ORGANIZATIONAL THEORY AND ADMINISTRATIVE BEHAVIOR

CHAIRMAN:

Richard Carlson, University of Oregon

PAPERS:

Authenticity, Role Distance and Organizational Climate: Toward a Conceptual Clarification

ROBERT B. BRUMBAUGH, West Virginia University

The organizational climates research by Halpin and Croft has produced an array of potentially fruitful concepts for organizational analysis. Among these is the construct of authenticity, a factor so-named which emerged from a Q analysis of their climates typology data. Use of this high-order intervening variable was invoked by Halpin and Croft because it helped to explain much of the variability within and among their subtests and climates findings.

Unfortunately, while appearing to possess conceptual promise for organizational research, subsequent empirical use of the authenticity construct has been precluded by lack of an appropriate operational definition for the variable. Although Halpin and Croft (and others) have proposed a number of approaches to the problem, none of these has been exploited to date by organizational researchers. Undoubtedly much of the apparent reluctance of researchers to tackle the task derives from the behavioral elusiveness of the prosaical definitions of the construct in the literature.

Sharing Biddle and Thomas' conviction that role theory concepts perhaps presently come closer to being the universal language of the behavioral sciences than any other vocabulary, this writer explores the possibility of linking Goffman's concept of role distance to Halpin and Croft's construct of authenticity in an effort to provide greater conceptual clarity to the latter and in order to suggest a means of facilitating its ultimate empirical operationalization.

Distinct similarities are noted in the authors' prosaical definitions of role distance and authenticity. For example, Goffman defines role distance variously as expressive denial of the virtual self that is implied in a role; those behaviors that are perceived by present others as suggestive of an actor's disaffection; and resistance against a role, or actions effectively conveying some detachment of the performer from a role he is performing. Halpin and Croft, on the other hand, have defined authenticity variously as professional roles remaining secondary to what the individual, himself, is as a human being, bringing one's own individual style to a role enactment, or, conversely, inauthenticity as being the use of one's role as a protective cloak to hide lack of personal identity, or performance of a role with an attendant denial of self. Other consistencies are compared in the prosaical definitions of role distance and authenticity.

Noting that, for purposes of operational definition, the behavioral vulnerability of Goffman's prosaical definitions of role distance are appreciably greater than are such definitions of authenticity by Halpin and Croft, specific recommendations for the operationalization of role distance are made and implications for further organizational theory and research are drawn.

◆ ◆ ◆

A Comparative Study of Organizational Decision Making in High Schools

WADE N. PATTERSON, Grants Pass Public Schools, Grants Pass, Oregon

The purpose of this empirical study was to determine if decision-making influence in high schools resides with persons occupying formal (official) or informal organizational positions. Two methodologies, reputational and issue analysis, treating five research questions, assessed influence in two schools selected under eight common criteria.

A series of various written questionnaires from several dozen respondents, analysis of formal organizational structure, and depth interviews with 35 persons provided the research data. Individual and group "lesser influentials" and "key influentials" were selected from the reputational technique, first on a general basis, then on a decision area basis, and finally, on the basis of specific decisions. The other approach included analysis of six specific issues selected by respondents in the schools. The role of persons and groups having influence was examined in each of the selected decisions.

Data supported 31 generalizations from one approach and seven, including the one following, from both. A form of elitism (as compared to pluralism) prevails in the schools. However, a relatively open decision-making climate serves somewhat to reduce this elitism. A relatively closed decision-making climate reinforces elitism and suppresses persons and groups that might otherwise influence decisions for the organizations.

Five greatly condensed conclusions, directly related to the research questions are 1) Organizational influence reputations correspond to formal positions when: respondents are aware of others' formal decision-making role; "lower' position holders don't exert influence; avenues of overt influence are closed. 2) Persons holding various bases of influence, by working through "higher" officials. 3) Influence is exercised by persons without authority positions under circumstances allowing the means listed immediately above. 4) Extra-organizational influence is manifested through confrontation of officials; personal influence of officials; reaching large numbers of people; authoritative sources; unanticipated chain events; planning and funding agencies; 5) Consistency, with exceptions, between reputations and perceived exercise of influence was noted.

◆ ◆ ◆

Administrator Competencies and Organizational Effectiveness

GORDON S. PURRINGTON, Florida State University, Tallahassee

The basic question of this research was organizational effectiveness. The aim of the study was to find out how supervisory skills facilitate or hinder system functioning. The four skills, human relations, conceptual, technical, and administrative, were so defined as to be related to the four functional needs of social systems, pattern maintenance and tension management, adaptation, goal attainment, and integration.

Two school systems in New York State were selected because of similar socioeconomic levels of the communities, and the vast differences in their "instructional quality," as measured by the N.Y.S. *Quality Measurement Study*. Data were obtained from the responses of teachers to a questionnaire, and from structured interviews with other organizational members.

The research design involved the comparison and contrast of the participating systems on each of the major variables through appropriate statistical techniques and analysis.

The findings indicated significant contrasts between the two systems. The supervisory competencies of administrators were stronger in more effective units. Additionally, a threshold effect appeared to be operating with regard to the human relations skill.

Teachers were more highly satisfied with administrators in more effective units. These administrators were more understanding of their teachers' viewpoints, tell teachers in advance about changes, and teachers are more sure of what their administrators think of them and their work.

Coordination in the more effective units was more general. Reliance was placed on the shared expectations and understandings among the organizational members. By contrast, the less effective units put heavy emphasis on regulations and operating procedures of the formal organization. Communication was also less personal; bulletins and memos were more often used. Complementing these findings, greater "singleness of direction" was found among the organizational members of the more effective units.

Significantly more tension and strain existed in the less effective system.

In short, the effective handling of the four functional problems confronting the school system (its organizational effectiveness) is related to the competencies of its administrators.

Relationships Among Dogmatism, Teacher Power, and the Organizational Climate of Schools

MARY M. BENTZEN and JERROLD NOVOTNEY, University of California, Los Angeles

The study was designed to investigate openness-closedness as it is conceptualized by three different instruments: Rokeach's Dogmatism Scale, the Organizational Climate Description Questionnaire (OCDQ) and a short instrument developed to measure teachers' power in decision-making within the schools. The Dogmatism Scale is based on the concept of open and closed cognitive systems in individuals. The OCDQ employs the idea of open and closed group structures. The teacher power measurement is concerned with one aspect of group organization which is theoretically related to open and closed school climates as they are described by the OCDQ.

Data obtained from these three instruments were available for three groups of elementary school teachers and principals: (1) One hundred parochial schools in the Los Angeles area; (2) Nine public schools in a small California city district; (3) Twenty public schools in twenty widely varied Southern California districts which belong to an organization dedicated to promoting educational change.

The unit of analysis for the study is the school. For some of the analyses, hypotheses were developed for a subset of the schools, and reexamined for the larger group. Specific analyses include the following:

1. Comparisons of principals' and teachers' mean dogmatism scores in schools with different climates on the OCDQ.

2. Relationships between dogmatism scores and the eight subtests of the OCDQ.

3. Relationships between dogmatism scores and particular items on the OCDQ.

4. The relationship of measured teacher power to OCDQ climate (using a simple ordinal scale developed from the OCDQ climate descriptions.)

Results of these analyses emphasize some of the difficulties in establishing a meaningful relationship between individual and group interpretations of the concept of openness-closedness. Results also help clarify the importance of the teacher power component in the measurement of open and closed school climates.

12:00 M to 12:55 P.M. **SESSION 11.4** **DIVISION D**

RESEARCH ON TESTING II

CHAIRMAN:

Jerome E. Doppelt, Psychological Corporation

PAPERS:

Effects of Item Order and Anxiety on Test Performance and Stress

RONALD K. HAMBLETON, The Ontario Institute for Studies in Education

The objectives of this research were to investigate the effect of item order on the performance of a test; on the amount of stress generated during a test; and on the performance of high and low anxious subjects.

106 high-school students completed the Achievement Anxiety Test. Two weeks later, they were randomly assigned to one of two treatment groups. Subjects in one group were administered a standardized mathematics achievement test with the items ordered from easy-to-difficult. Subjects in the second group took the same test with the order of items reversed. A physiological indicant of stress, heart-rate, was measured three times during the test using a pulsometer. The three heart-rate measures for each subject were averaged to obtain a stress score.

Results of this study confirmed the finding of other researchers that the mean number of correct answers for the test questions arranged in the difficult-to-easy order were significantly lower than the mean number for the questions arranged in the reverse order. In addition, it was found that the average stress score of subjects was significantly higher under the difficult-to-easy order; this indicates that item order does effect the amount of stress induced in a test situation.

Finally, contrary to expectation, the test performance of high anxious subjects did not differ from the low anxious subjects under the more stressful difficult-to-easy order of items. However, low anxious subjects did significantly better on the less stressful easy-to-difficult order.

◆ ◆ ◆

An Experiment Involving a Probability Measurement Procedure

THOMAS ROMBERG and JACK SHEPLER*, University of Wisconsin

Multiple choice tests comprised of difficult items produce low reliabilities using standard test taking and scoring procedures. In general, tests measuring problem solving, insightful, or creative cognitive behaviors have these characteristics. The purpose of this study was to see if a non-standard test taking and scoring procedure would provide useful, reliable information on such a test.

A test of 15 difficult items was used in this study. The test was constructed from a variety of items designed to measure "Insightful Mathematical Abilities" as a part of the achievement tests used in the National Longitudinal Study of Mathematical Abilities.

The experimental procedure, a probability measurement procedure called PMP, is designed to be sensitive to a subject's partial knowledge about an item. In this procedure the student is presented a multiple-choice item with five alternatives and asked to specify what he believes to be the probability of correctness of each alternative, where the total of the probabilities for the five choices is 1. The student receives a score on any item equal to the probability he assigns to the correct choice. The student is instructed that he can maximize his score if he honestly reflects his degree-of-belief probabilities as to the correctness of each of the choices for an item.

The experiment involved two randomly selected groups of twelfth grade mathematics students. One group was taught PMP and told to use this procedure on the test. The second group was given the standard test taking instructions when taking the test.

Data from the two groups were used to compare test reliabilities, test means, and item characteristics. The results of the experiment were: (1) the reliability coefficient on the test for the PMP group was greater than that for the standard group, and (2) the mean score obtained by PMP scoring was greater than the mean score for the PMP group scored in the standard way which in turn was greater than the mean of the standard group.

From the results of the experiment one is led to believe that the PMP training introduced a test-taking style which changed the performance of the PMP from that of the standard group. This test taking procedure, used with a multiple choice test made up of difficult items, produced more useful information than standard test taking procedures.

◆ ◆ ◆

Partial Control of Guessing on Multiple-Choice Tests by Weighting Systems

C. E. ALEXAKOS and C. E. SMITH*, West Virginia University

The frequently used formula for corrections for guessing on multiple-choice tests is given by (1.1), where, Corrected $X_i = R_i - \frac{Wi}{m\text{-}1}$ Formula

(1.1) is not universally accepted by psychometrists as the best method of correcting for guessing (Nunnally, 1967). The present study was designed to control partially the effect of guessing by assigning weights to the alternative choices contained in each item.

Two weighting systems were contrasted with the binary method of scoring multiple-choice tests (0 for failing and 1 for passing an item). According to the first weighting system, testees were instructed to assign weights (w) of 3, 2, 1, and 0 to the alternatives of a 50 item test, the items of which contained 4 alternative choices. According to the second, the same weights were assigned to each alternative by the instructors of a course in Educational Measurement. Two *randomly* equivalent forms (A and B) of the same test were used. The test items were intended to evaluate the students' achievement in the course; therefore, the two forms could be considered as equivalent forms of an achievement test.

Items were scored in three ways: (1) *Binary Methods of Scoring.* If a student weighted the keyed correct alternative with a "3", he received one point for the item; if he assigned any other weight to the correct alternative, he received no credit; (2) *Internal Weighting System of Scoring.* The student was credited with the w_i he had placed on the correct alternative; (3) *External Weighting System of Scoring.* The student received the w_i predetermined by the instructors which corresponded to the alternative he thought to be most correct; i.e., the alternative he had weighted with a "3".

In a sample of 78 subjects, the results of the study indicated that the internal system of weights resulted in significantly higher scores than the adjusted binary scores for both Forms A and B at the .01 level. Also, the internally weighted scores were superior to the externally weighted scores at the .01 level for Form A, but no significant difference was observed between internal and external scores for Form B. The externally weighted scores were higher than the binary at the .01 level of significance for Form A and at the .05 level for Form B. In order of efficiency the three systems were ranked as follows: (1) Internal weights, (2) External weights, and (3) Binary scoring.

Form A proved more difficult than B despite the random assignment of items to the two forms. From a pool of 180 items, 50 were assigned randomly to Form A and 50 to Form B. No ready explanation exists for this phenomenon. It was observed, however, that the first item in A was a difficult one, while the first item of Form B was an easy item. This fact may support Cronbach's (1946) theory of "sets" or attitudes toward test-taking.

From the results of this study it was concluded that weighted scoring can be somewhat more efficient in controlling the effect of guessing on multiple-choice tests than the usual binary method of scoring. Ranking of the alternatives of an item by the testees seemed to be the best of the three methods of scoring with regard to the control of guessing. It was also observed that the difficulty of a test may be related to the degree to which subjects guess.

◆ ◆ ◆

A Model for the Evaluation of a Testing Program

NANCY JORDAN UNKS and RICHARD C. COX, University of Pittsburgh

The evaluation of an educational program typically implies measurement; measurement, in turn, implies testing in one form or another. In order to carry out the testing necessary for the evaluation of an educational program, researchers often develop a complete testing sub-program. The evaluation of the total project, therefore, may depend upon the testing sub-program. If the testing program is something less than adequate the evaluation of the total project may be suspect. The point is that researchers should pay as much attention to the evaluation of a testing sub-program as they do to the evaluation of the total project of which it is an integral part.

The proposed model for evaluating a testing sub-program includes the following steps: (1) Define the unique objectives of the testing program. These are generally subordinate to total project objectives and provide for obtaining information about pupils, curriculum and measuring instruments, as well as the functioning of both the testing sub-program and the entire project. (2) Define the testing program with regard to personnel and facilities, planned and actual functions, and products. (3) Plan and carry out evaluation of the testing program concurrent and consistent with the total program evaluation. This would include objectively assessing the achievement of the testing program objectives and observing, perhaps subjectively, any unplanned results of the testing program. (4) Attach a 'valuation' to the testing program to answer the question, "Can an evaluation of the total project based upon this testing program be considered sound?"

In order to clarify, the model, its application to a specific testing program will be described. This example is the testing program for the Individually Prescribed Instruction Project of the University of Pittsburgh's Learning Research and Development Center.

12:00 M to 12:55 P.M. **SESSION 11.5** **DIVISION C**

CONCEPT LEARNING II

CHAIRMAN:

Robert A. Rosemier, Northern Illinois University

PAPERS:

The Effect of Instruction and Concomitant Variables on Multiple Categorization Ability

JOSEPH C. EDWARDS, Southwest Regional Laboratory for Educational Research and Development, and EVAN R. KEISLAR, University of California, Los Angeles

This investigation was conducted to determine if first grade boys' multiple categorization behavior could be modified by means of group instruction. It was hypothesized that instruction would facilitate multiple categorization performance and that one method, T_1, of the overt presentation of concept dimensions, would be more effective than T_2 which consisted of an inferential or "guided discovery" approach. In addition, four other variables, intellectual ability, delay of immediate gratification, task-oriented dependence, and response latency were used as classification factors to ascertain whether they were related to the subjects' ability to categorize objects according to several attributes or relations.

The total sample consisted of 154 first grade boys from greater Los Angeles elementary schools. Subjects were assigned to one of three groups, T_1, T_2, or control. Both treatment groups received instruction in matching the same common objects. T_1 was given a verbal cue describing the concept dimension to be used as the basis for matching two different objects. T_2 was only provided with the label for the stimulus object and had to infer the basis for matching from the feedback provided by the experimenter.

To assess multiple categorization performance, a test was administered following two days of instruction. The test consisted of eight depicted objects, each of which could be classified with eight or nine other common objects such as pictures of furniture, animals, vehicles, people, and the like. There was a total of seventy-one possible items, responses to which were grouped into one of three concept classes, analytic, relational, or inferential-categorical.

The treatment of the data consisted of performing univariate and multivariate analyses of variance on total multiple categorization scores as well as the concept class scores obtained from the criterion test.

Since subjects in both instructional treatment conditions made significantly more categorizations than the controls, the findings indicate that children's categorizing skills can be shaped and modified on a group instructional basis. Second, particular instructional techniques may be more appropriate than others for teaching certain classes of concepts; T_2, or the inferential method, taught the subjects to categorize objects using relational concepts better than T_1. And third, certain variables, namely intellectual ability and response latency, are related not only to a child's ability to categorize objects in his immediate environment, but also to the particular category of concepts or schema he employs in this classification process.

The results obtained will be discussed in terms of their relationship to styles of conceptualization or concept class preferences and their implications for educational practices in the everyday classroom.

◆ ◆ ◆

The Effect of Different Verbal Directions upon the Interpretation of Conceptual Rules by Middle and Lower Class Kindergarten Children

SALLY A. THOMAS, SAMUEL R. SCHUTZ, and EVAN R. KEISLAR, University of California Los Angeles

The purpose of the study was to assess the abilities of 60 middle and lower class kindergartners to interpret four conceptual rules (affirmation, negation, conjunction, disjunction) when the directions were given in four different sentence formats (keywords-incomplete sentences, and declarative, interrogative, imperative sentences).

Each subject was given four 15 minute test sessions of 40 items each on four successive days. Each item consisted of the presentation of a 6 x 8 card containing 2 pictures and a tape recorded direction designating the correct conceptual rule. Subjects were required to point to the picture which matched the rule.

The data were analyzed using a Latin Squares (repeated measures) design. Comparing conceptual rules, sentence format, and social class the following results were obtained:

1. For the total test the difference in performance of middle and lower class children was not reliable.

2. Performance on conceptual rules differed significantly with affirmation by far the easiest, and negation the most difficult, for all subjects.

3. There was a significant difference attributable to sentence format of the directions. Some formats were much easier for the children than others.

4. Additionally, the interaction between social class and sentence format was significant at the .05 level. The differences took the following form. Middle class kindergartners did relatively better when the sentence format was interrogative; whereas lower class children did better if only key words in incomplete sentences were given.

One interpretation hypothesized for such a finding was that lower class children, to whom the use of function words is little known, are not aided by their inclusion in full sentences, but perform better with key words in incomplete sentences. Further interpretations are related to other findings on differences in language patterns used by different socioeconomic groups.

◆ ◆ ◆

Unidimensional and Multidimensional Strategies in Concept Identification

J. A. EMRICK and M. C. WITTROCK, University of California, Los Angeles

In this experiment, three strategies were investigated for teaching young children to transfer scaled relationships among stimuli. The strategies were 1) Within, *W,* where each problem presented five variations within a single dimension; 2) Across, *A,* where each problem presented variations across each of four dimensions; and 3) Across-Within, *AW,* where each problem presented all variations across and within all dimensions. The first two strategies are considered unidimensional, the third multi-dimensional. Rules presented during training described the scaled relationships within and across the dimensions for the original stimulus matrix ($5 \times 5 \times 5 \times 2$ elements), but not for the transfer matrix ($5 \times 5 \times 5$ elements).

A random sample of 64 second grade *S*s was drawn from a population of over 200 second graders. This random sample of 64 children was assigned individually at random to the experimental and control treatments in blocks according to IQ level, generating a randomized black design. Automated equipment was employed for stimulus presentation and data collection during training and testing. All *S*s received familiarization training. The experimental groups received 120 training problems, the last 60 problems contained original and transfer stimuli. The control groups received no training problems. All groups were then given tests over training problems, near transfer problems, and remote transfer problems.

Analyses of variance and multiple comparisons tests indicated that the means of the *W* and *AW* groups were highest (p. $<.001$) on a test over the training problems, but only the *AW* group mean was superior on near transfer ($p < .05$) and remote transfer ($p < .01$)tests. The data are discussed in terms of a theory of instruction emphasizing verbally mediated transfer.

◆ ◆ ◆

The Effect of Subject-Determined Verbalization on Discrimination Learning in Preschoolers

J. L. WOLFF, University of Illinois

An experiment was conducted to examine Reese's mediational deficiency hypothesis which states that verbal responses, though available, do not readily mediate discrimination learning in very young children.

All Ss performed an initial form discrimination with stimuli varying in a single dimension (i.e., form) and were then nonreversal-shifted to either a brightness or a size discrimination with stimuli varying simultaneously in the dimensions of brightness and size (form held constant). Half of the *S*s verbalized their choices during both the original-learning and transfer-learning periods, while the remaining half responded throughout the entire task without speaking. On the basis of previous results, it was predicted that requiring subjects to "name" the stimulus to which they wished to respond before making motor choice would cause most *S*s to use brightness labels to describe the stimuli during transfer learning regardless of the dimension relevant. Contrary to the Reese's hypothesis, it was further predicted that verbalization would retard learning of a size discrimination and facilitate learning of a brightness discrimination as compared to a non-verbalization control.

The principal finding was that during transfer learning, verbalization facilitated the acquisition of a brightness discrimination ($p < .01$) but had no effect on a discrimination of size. Other results were that (a) size was learned more rapidly than brightness, and (b) as expected, brightness labels were used to describe the stimuli regardless of whether the relevant dimension was size or brightness.

These results were interpreted as supporting neither Reese's hypothesis nor an alternative hypothesis proposed by Kendler. It was finally concluded that the results of the study were best explained by a position recently developed by White.

12:00 M to 12:55 P.M. **SESSION 11.7** **DIVISION D**

CONTRIBUTED PAPERS I

CHAIRMAN:

Robert D. North, Psychological Corporation

PAPERS:

The Relations Between the Distributions of Course Grades and Grade Point Averages

WILLIAM A. MEHRENS, Michigan State University and BRUCE G. ROGERS, Arizona State College

Many measurement specialists recommend that the faculty of an educational institution establish a uniform classroom grade distribution policy, with an approximate percentage considered appropriate for each letter grade, and encourage general observance of it through the entire staff. They contend that this practice makes course grades have a more common meaning across teachers and subjects. Most schools also have minimum grade point average (GPA) restrictions for participation in various activities. Unfortunately, the determination of the recommended classroom grade distributions and the establishment of minimum GPA requirements are generally done independently, with little consideration of the interrelationship between the two problems. The present investigation was initiated to explore the interrelationship between the distributions of course grades and grade point averages. A simple mathematical model depicting the relationship was proposed and empirically tested.

The model, obtained from a fundamental theorem of multivariate analysis, simply asserts that the GPA distribution is completely determined given normally distributed grade distributions (with equal variances) and the intercorrelations of the grade distributions. Since the assumptions of normal grade distributions with equal variances are never precisely met, the robustness of the model was tested empirically. Grades were obtained for a sample of 837 undergraduate students who had completed 12 terms of four basic courses. The data for these students were used in the model and the predicted GPA distribution was compared with the GPA distribution. The analysis pointed to the conclusion that the variances and intercorrelations of the classroom grade distributions are sufficient to determine the resulting variance and shape of the GPA distribution within reasonable accuracy.

Hypothetical examples were shown to indicate the importance of the intercorrelations in the model. It was suggested that those who are setting policies concerning minimum GPA requirements for various activities should look at the intercorrelations as well as the typical grade distributions of the relevant courses.

A Multivariate Study of a Set of Educational Objectives

ROBERT M. PRUZEK, State University of New York at Albany

The present paper is concerned with the development and evaluation of two multivariate techniques for studying two kinds of student judgments for 24 educational objectives. These objectives were constructed to be generally appropriate for an introductory course in educational research, taught during the summer to 89 students. Similarity judgments and preference judgments were obtained at both the beginning and the end of the course.

The objectives were initially devised to fit a 2^3 balanced factorial design; thus, eight *generally* distinct categories of objectives could be identified according to combinations of the following factor levels: (1) strong verb *vs* weak verb, (2) test development *vs* statistics and research design and 3) theoretical *vs* applied situation.

For the similarities study, students independently sorted objectives into mutually exclusive and exhaustive categories according to perceived similarity. Resulting categories were analyzed using latent partition analysis (LPA, Wiley 1967)) and 8 clearly interpretable categories were derived and interpreted for both the pre- and post-course data. Similarities were found between these two sets of categories, but neither of these categorizations corresponded closely with the initial design categories.

For the preferences study, students ranked the objectives according to relative preference using a Likert-type scale. These data were used to generate correlation matrices which were then analyzed using principal components analysis (PC) followed by normal varimax rotation. Eight interpretable dimensions for preference were derived for both the pre- and post-course data. As with the LPA study, substantial similarities existed between these sets of derived components but neither set was closely associated with the initial design categories.

Results of the individual analyses and relationships among them are discussed in detail, including substantive interpretations. Further, speculations are made as to expected relationships between the PC analysis of preference data and a possible structural covariance analysis (Bock and Bargmann, 1966); thus, association between a type of factor analytic study and a type of analysis of variance study is considered.

◆ ◆ ◆

Prediction of Effects with Selected Characteristics of Linear Programmed Instruction

MARTIN E. SMITH, Bell Telephone Laboratories, Incorporated, and
WARREN F. SEIBERT, Purdue University

The principal purpose was identification of minute, often linguistic features of instructional materials which correlate with observed instructional effects. The thesis was that such correlations identify stimulus variables meriting more precise study. Multiple regression analysis with double cross-validation was employed to estimate the effects prediction that is possible with knowledge of stimulus features. Correlations among features were factor analyzed and factor scores then employed as predictors of effects.

Paid volunteer undergraduates were assigned randomly to an instructed experimental group, an uninstructed control, and a "cloze score" group, with *n*s, after modest attrition, of 61, 60, and 14, respectively. Over five weeks, experimentals studied Holland and Skinner's *Analysis of Behavior,* a linear teaching program with about 2200 frames. Criterion items were administered at four points in this period and again, *in toto,* three weeks after instruction concluded. Controls served only to take the criterion test, thus, in estimating items' pre-instructional difficulty levels. "Cloze score" *S*s worked through a modified Holland-Skinner program, one from which every fifth word was deleted. Their task was to supply deleted words; indirectly, they provided readability estimates for sequences. Thirty-six other features (predictors) described each sequence. Six indexes of instructional effect represented absolute, relative, and residual gain for "immediate" and "delayed" posttest items.

Success in identification of correlated features and multiple prediction of effects was initially good, but reduced severely in cross-validation. Stability of prediction improved with factor scores as predictors. Factor analysis of features yielded nine orthogonal factors; eight were named.

◆ ◆ ◆

Empirical Evidence and Personal Consistency

TOM RUSK VICKERY, Northwestern University, and BOB BURTON BROWN, University of Florida

The purpose of this research was to investigate the use of measures of beliefs in conjunction with systematic observations of classroom behavior as a means of promoting consistency between beliefs and practice. The *Personal Beliefs Inventory* (PBI) and the *Teacher Practices Inventory* (TPI) were used to measure personal and educational beliefs, respectively, and Rokeach's *Dogmatism Scale* (D Scale) was used to measure openness to change. The *Teacher Practices Observation Record* (TPOR) was used to describe classroom behavior. The PBI, TPI, and TPOR measure beliefs and behavior along a common dimension—in terms of agreement-disagreement with John Dewey's philosophy of experimentation.

The data obtained by these instruments were used in an extended in-service program as the basis for a discussion of the relationship between belief and practice. Utmost care was taken to avoid pressing for or against any belief or behavior; instead the teachers were urged to believe and behave consistently, regardless of what beliefs they held.

Posttest administration of these instruments revealed (1) that the subjects did not become more consistent as measured by the instruments and (2) that both before and after the in-service program there were racial differences in belief, in behavior, and in openness to change.

1:00 to 1:55 P.M. **SESSION 12.5** **DIVISION B**

LANGUAGE INSTRUCTION FOR DISADVANTAGED STUDENTS

CHAIRMAN:

Elliott W. Eisner, Stanford University

PAPERS:

The Effects of Different Language Instruction on the Use of Attributes by Pre-Kindergarten Disadvantaged Children.

DAVID E. DAY, Emory University

Sigel's system for organizing language grouping preference behavior was used in analyzing attributes used by pre-kindergarten disadvantaged children. The purpose was to determine the similarities and differences of categorization style resulting from two different methods of instructing the children in the use of language as an intellectual tool.

Group A (N = 49) consisted of four randomly selected pre-kindergarten classes. The children were organized for instruction in groups no larger than seven, taught by a lead, an assistant teacher, or an aide. Curriculum and teacher behavior were specified in detail following the style of Carl Bereiter.

Group B (N = 38) consisted of four randomly selected pre-kindergarten units. These children were organized on the basis of a developmental units-of-work approach. Each unit was taught as a group by the lead teacher.

At the end of ten months' instruction, all subjects were individually asked to tell all they could about a toy turtle, a cup, and two toy cars. An analysis of the recorded responses revealed the following:

1. There were no significant differences between the groups on total language production.

2. Group B produced significantly more relational-contextual words describing the function of objects) than Group A.

3. Group A produced a significantly greater number of categorical-inferential or concept words than Group B.

4. There were no significant differences between groups in the total number of descriptive part-whole words used.

5. Descriptive part-whole words were further analyzed according to: (1) Numbers and kind or use of numbers, adjectives, or adverbs to describe parts of the objects. (2) Color and form in which the subjects used color, shape, or character of an object to describe it. In each case Group A produced significantly more complex descriptive words.

Group A was thus shown to be able to use language in a clearer and more economical way than Group B. Children in Group B describe objects more by their function than did Group A. A highly structured semi-conditioning based approach to instruction did not impair fluency. Group A children gave evidence of having transferred behavior from the teaching situation to the interview. For these disadvantaged children with severe language deficits the procedures used for Group A were more effective than those employed with Group B.

◆ ◆ ◆

The Language of the Inner City Child: A Comparison of Puerto Rican and Negro Third Grade Girls

MARCIENE S. MATTLEMAN, and ROBERT EMANS, Temple University

This investigation compared the oral language of two groups of inner-city girls, Negro and Puerto Rican. It was an attempt to study in depth a small number of children from both quantitative and descriptive aspects.

The sample was composed of five Negro and six Puerto Rican girls from the third grade who were identified as highly verbal in English on the basis of teacher judgement. Each child was asked to respond individually in an interview to a series of three pictures from the *Daily Language Facility Test.* Each of the resulting oral language protocols was analyzed for facility as measured by the test itself, for fluency as determined by word count, and for syntactic structure using a category system developed by one of the investigators.

The results of this study indicated that the Negro children tended to score higher ($P < .06$) on the *Daily Language Facility Test* and were more fluent ($P < .01$) than Puerto Rican children. An analysis of the syntactic structure showed that Negro children used more prepositional phrases, more transformations, greater variety in sentence pattern, and more linking verb construction than Puerto Rican children—suggesting more sophisticated speech.

The results of this investigation indicate replication is warranted. The Puerto Rican children in this study appeared to possess less language proficiency than Negro children. If the results of this study are supported by additional research, Puerto Rican children would seem to require a different type of instructional situation or perhaps curricular objective (e.g. reading) than Negro children at the third grade level.

◆ ◆ ◆

A Model for Teaching Standard English to Non-Standard English Speakers

LLOYD LEAVERTON and MILDRED R. GLADNEY, Chicago Board of Education

The purpose of this study was to investigate the effectiveness of an oral language program developed by the authors. The program was developed for those urban Negro children whose speech differs from standard English.

However, if the model proves effective with this group, it should also be applicable to other groups of children whose dialect differs from standard English.

Development of Materials. In developing this approach, conversations with primary children in two ghetto schools were tape recorded. The tapes revealed striking differences from standard English in the areas of copulative verbs and subject-verb agreement. A program was therefore developed focusing on this area of oral speech. The program consists of activities and materials based on three assumptions:

1. The learning sequence should start with an actual statement made by the child.

2. The transition from the child's dialect to standard English should be additive and not involve unlearning. For example, "My mama pretty.", can be restated in standard English by adding *is.*

3. The teacher should concentrate on only one variable at a time.

Collection of data. Samples of informal oral speech were tape recorded from all children in both the experimental and control groups. The recordings were analyzed with respect to the extent of correspondence to standard English, or to the non-standard dialect when any one of the six verb forms comprising the experimental treatment occurred in their speech.

Each child was rated as a plus or minus with respect to standard English usage for each of the six verb forms. The X^2 technique was then used to determine the significance of the difference between the groups.

With the exception of the conditional verb form, the findings show positive trends, but are generally inconclusive.

Extensive investigations are needed to determine why the experimental treatment appeared more effective with some verb forms than with others.

Perhaps the most important implication of the study was with respect to the concept of "everyday talk" and "school talk" and its influence on the attitude and behavior of the teacher toward the children's oral speech.

◆ ◆ ◆

Year II Findings for the San Antonio Language Research Project

RICHARD D. ARNOLD, The University of Texas at Austin

The primary purpose of this study is to determine whether a specially designed oral language program centered around science materials was effective with disadvantaged Mexican-American children learning English as a second language.

The two samples consisted of children from a densely populated urban area where the median family income was less than $3,000. Eighty-one classes and over 2,000 children were included in the second year study. The four treatments were:

1. *Oral-Aural English (OAE).* Experimental language program taught in English using AAAS science materials.

2. *Oral-Aural Spanish (OAS).* Experimental language program taught in Spanish using AAAS science materials.

3. *No Oral-Aural (NOA).* AAAS science materials taught in English without a special language program.

4. *Control.* Science and Language Arts as prescribed by the district curriculum guide, taught in English.

English and Spanish reading measures were administered to first and second grade children as both pre- and post-tests. Intelligence tests were also administered. Analyses of covariance were used. Spring test scores were the criterion; covariates were fall test scores and intelligence test scores.

Analyses of the second grade data revealed few significant differences between the OAE and OAS groups, except in word knowledge, which favored the group taught in Spanish. The NOA treatment was found to excel both oral language treatments on virtually every reading test used as the criterion. Analyses of the first grade data are not quite complete, but there appear to be few significant differences. Definitive conclusions will be available after more careful examination of the findings.

Recognizing certain limitations in the study, it appears that analyses of reading achievement scores at first and second grade have failed to provide evidence of the beneficial treatment effects previously detected when oral language tests were used as the criterion.

1:00 to 1:55 P.M. SESSION 12.6 DIVISION E

PSYCHO-SOCIAL DEVELOPMENT

CHAIRMAN:

Murray Tondow, Palo Alto Unified School District

PAPERS:

The Perceptions of School Personnel and Parents, and Mental Hygienists Toward Behavior Problems of Children

GERALD G. MANSERGH, University of Michigan

This study was developed to assist in the implementation and evaluation of an exemplary school mental health program involving the use of mental hygienists in helping school personnel and parents develop more positive attitudes toward children.

A modified version of the Wickman Behavior Rating Scale was used to obtain the perceptions of school personnel, parents, and mental hygienists toward the importance of certain "inferred" behaviors in the future adjustment of elementary and/or secondary age children.

Within a five school district area of suburban Minneapolis, the instrument was administered to all elementary principals (63), all elementary school psychologists (10), a random sample of elementary teachers (137), a random sample of parents of elementary children (111), all secondary principals (19), all secondary assistant principals (24), all secondary counselors (76), a random sample of secondary teachers (127), and a random sample of parents of secondary youth (71).

In addition, a mental hygienist criterion group of psychiatrists (10), clinical psychologists (6), and psychiatric social workers (22) from the six major mental health clinics in the Twin Cities metropolitan area were administered both the elementary and secondary forms of the instrument.

Utilizing a multiple linear regression computer program, "inferred" behavior ratings of the various groups of school personnel and parents at the elementary level were compared one to another and to the criterion group of mental hygienists; similar comparisons were made with the ratings of secondary level respondents. A level of .05 was considered a significant difference.

General findings were as follows:

1. School psychologists and counselors were closest to mental hygienists in their perceptions; all were more concerned with recessive behaviors.

2. An extreme number of significant differences exist between the perceptions of elementary principals or parents and school psychologists. The former groups are significantly more concerned with aggressive behaviors.

3. At the secondary level, mental hygienists and parents or teachers showed the greatest numbers of significant perceptual differences.

4. Among secondary respondents, the greatest number of significant differences in ratings was found with the counselor-parent comparison.

◆ ◆ ◆

Classroom Behavior, School Achievement, and Social Adjustment: A Longitudinal Study

JOHN F. FELDHUSEN, Purdue University, and JOHN R. THURSTON and JAMES J. BENNING, Wisconsin State University-Eau Claire

Do children who are aggressive and disruptive in school achieve at lower levels, and is their social adjustment less adequate than children who behave in socially acceptable ways? A total of 1109 children were nominated by third and sixth grade teachers as persistently displaying aggressive and disruptive (disapproved) or socially acceptable (approved) behavior. Four years later, a sample of 160 was drawn, divided equally by behavior status (approved or disapproved), grade (7 or 10), and sex for a longitudinal study of the long-range effects of classroom behavior on standardized achievement test performance (STEP), teacher grades, and teacher ratings of social adjustment. An analysis of covariance design (anacova) with IQ as the covariate, was used in the analyses involving STEP scores and teacher grades; and an analysis of variance design (anova) was used for the social adjustment scores. In both the anova and the anacovas, the independent variables were behavior, grade, and sex while the dependent variables were the achievement and adjustment scores.

It was found that after four years all achievement and adjustment means of the children who had been nominated as aggressive and disruptive were significantly lower than the means of their classmates who had been nominated as displaying socially approved behavior.

The results are discussed in relation to other findings of the Eau Claire County Youth Study, a long term research of classroom behavior.

◆ ◆ ◆

Stability of Deviant Behavior Through Time

HARVEY CLARIZO, Michigan State University (T. Antoinette Ryan, Oregon State University)

Clinicians and developmental psychologists have traditionally differed in their answers to the question of stability of deviant behavior in children. The former are more inclined to view childhood problems as being of a chronic nature, whereas the latter view them as being of a nonpersistent nature.

Illustrative of the major findings of this review are the following: (1) It appears that certain "types" of disturbed children do contribute more than their share to the population of adult psychiatric disability. The extent of overlap between childhood and adult disturbance remains, however, uncertain. (2) While it is commonly assumed that adult behaviors and personality are established in early life experiences, and while there is a body of empirical research demonstrating the stability of personality over time, there is nonetheless a genuine danger in over-generalization. (3) Somewhat contrary to prevailing clinical belief, it is aggressive, antisocial, acting-out behavior of a severe nature that is most predictive of later significant adult disturbance and most deserving of our treatment efforts. (4) Although two of three unsocialized aggressive youngsters continue in their antisocial ways, it must not be assumed that the majority of young norm violators become adult criminals. (5) Shyness or withdrawn behavior tends to disappear with advancing age. (6) Neurotic symptoms (fears, hypersensitiveness, tics, etc.) often presumed to be the precursors of adult neurotic disturbances have also been found lacking in prognostic power. (7) The probabilities of remission characterizing childhood schizophrenia and infantile autism are indeed low. (8) Generally speaking, it is very difficult to postulate any direct causal relationships between early childhood maladjustments and later specific psychiatric disability.

In our present state of knowledge, we can conclude that there is at best only a mild or moderate evidence to support the notion that disturbed children turn into disturbed adults. Since the less noxious childhood disorders, e. g., shyness, are more common than the more severe disorders, e.g., childhood schizophrenia, it would seem that, all in all, change appears to characterize the curse of behavior deviations in children as much as or more than chronicity or stability. The concept of emotional disturbance in children as a progressively deteriorating condition is thus called into question.

◆ ◆ ◆

Socio-linguistic Dimensions of Youth Culture

EDWARD A. NELSEN and EDWARD ROSENBAUM, University of Wisconsin

The investigation was designed to examine a word association paradigm for analyzing subcultural groups in terms of language contents. The paradigm was applied in a developmental analysis of slang usage by adolescent groups. Presumably, the language contents peculiar to particular age groups reflect the basic concerns, motives, and interests of the members of the groups. Moreover, changes in language contents according to age may provide evidence concerning the role of the peer groups across age.

Aproximately 600 junior and senior high school students were asked to list the slang words they used in talking among themselves about cars, money, cigarettes, alcohol, boys, girls, popular persons, unpopular persons, clothes, and appearance. *S*s were arranged into groups of four, uniform according to age and sex. Five minutes were allowed for each group to list all the words on a given topic.

Counting the number of responses to each stimulus-topic at each grade level indicated, as one might expect, that amount of slang knowledge increased with successive grade levels. The increase in number of words listed, however, was not uniform for the various topics, and boys and girls differed considerably in some respects. Examination of Table 1 shows that for boys the largest increase in number of words for "an unpopular person" came during 9th and 10th grades; for "girls", during 10th grade; for "boys", during 8th grade; for "cars", during 11th grade; for "alcohol", during 12th grade.

For girls, the largest increases in words for "an unpopular person" came during 8th and 10th grades; for "girls", during 10th grade; for "boys", during 10th grade; for "clothes, styles and appearance," during 10th grade; and for "alcohol," during 12th grade.

These findings imply that these various topics are of particular concern during these various developmental periods. In the case of cars and alcohol, the increased concern appears to coincide with changes in age-related legal prohibitions and sanctions relating to these activities. Although the findings from this one sample are merely suggestive, they do seem to indicate that a method of measuring slang word associations is a promising means of assessing youth culture based concerns, interests, and motives as reflected by the language used by teen-agers.

1:00 to 1:55 P.M. **SESSION 12.7** **DIVISION C**

SOCIOECONOMIC STATUS II

CHAIRMAN:

Lauren B. Resnick, University of Pittsburgh

PAPERS:

The Auditory Memory of Children from Different Socio-Economic Backgrounds: A Partial Replication

LOREN S. BARRITT, University of Michigan

It was the purpose of this study to examine the differences in auditory memory capacity of kindergarten, 1st, and 2nd grade children from lower and middle socio-economic backgrounds.

Auditory memory capacity for materials at 4 levels of linguistic sophistication was tested. These levels were:

1. nonsense trigrams;
2. high frequency nouns;
3. anomolous sentences;
4. meaningful sentences.

Subjects were 17 children each from the kindergarten, 1st, and 2nd grades at 2 schools. One school enrolled the children of professional, white-collar, university-educated parents (middle socio-economic), while the other school located near a large city enrolled the children of blue-collar workers who had not had the same educational opportunities (lower socio-economic).

At each level, materials consisted of lists of increasing length. Subjects were tested individually. Lists were presented on tape, and the entire testing period was recorded. A child's score at each level was the number of items correctly recalled in order on the last correctly recalled list prior to failure on the next two lists.

A three way ANOVA was performed on the scores for groups (lower socio-economic and middle socio-economic) levels (1 thru 4) and grade (K, 1, and 2).

All main effects of the ANOVA were significant. The interaction between groups and task levels was significant as was the interaction between grades and levels. Examination of the means indicates there is an increasing capacity to recall information as tasks permit the use of prior language habits. There are larger discrepancies between the memory capacity for sentences of lower and middle socio-economic kindergarten and first grade children than for their second grade counterparts.

In a previous study (Barritt et al., 1967) a remarkable similarity in the auditory memory capacity of children from different socio-economic environments was found. The present study was an attempt to replicate that research with groups of children from even more divergent socio-economic backgrounds. The present study indicates that memory capacity for materials with different degrees of linguistic relevance are different for "advantaged" and "disadvantaged" groups of children.

◆ ◆ ◆

Immediate Memory of Three Word Classes as a Function of Social Class and Grade

SAMUEL R. SCHUTZ and EVAN R. KEISLAR, University of California, Los Angeles

Some investigators have proposed that underprivileged young children suffer particular deficiency in the use of certain word classes. The suggestion is advanced that these children have the least trouble with nouns, moderate difficulty with verbs, and most trouble with functors (prepositions, articles, etc.). This study investigated the relation of social class (low and middle) and grade level (kindergarten, first, second) to the ability of children immediately to recall words from the three different word classes. One school from each social class was represented, with thirty Ss per school and ten Ss per grade.

A sample of eight one-syllable words from each class was selected from the 500 most frequent words in the original Thorndike count, and arranged in pairs, triads, etc., by word class. Ss were tested individually using standard "digit span" procedure. One trial was given per item; testing was discontinued for each word class when an error was made in that class. Ss score for each word class was total number of words correctly recalled.

A three way analysis of variance indicated that middle class children were superior to lower class, and each grade performed better than the grades below it. Middle class children performed better in kindergarten and grade one for all word classes; but, at grade two, this difference applied only to functors. At each grade, an ascending order of difficulty was obtained for nouns, verbs, and functors. The results support the findings of other investigators that lower class children are most deficient with functors, and further suggest the possibility of developing a new type of measure for assessing language deficiency.

◆ ◆ ◆

Socioeconomic Status and Learning Proficiency in Young Children

WILLIAM D. ROHWER, JR., NANCY SUZUKI, and LINNEA EHRI, University of California, Berkeley

Previously it was reported that lower-SES elementary school children can learn a pictorial paired-associates (PA) task as efficiently as middle-SES children, despite the fact that the performance of the former is markedly inferior to that of the latter on standardized tests of school achievement. The present experiment was designed to specify the conditions under which the PA performance of children from the two populations would parallel their achievement-test performance. The initiating hypothesis was that the lower the degree of environmental support for the use of mnemonic elaboration in the PA task, the greater the likelihood that middle-SES children would perform better than lower-SES children.

Of a total sample of 240 children, 40 were drawn from each of the following six populations: Middle-SES kindergarten, first, and third grades; and lower-SES kindergarten, first, and third grades. Within each subsample, 10 *S*s were assigned randomly to each of four experimental conditions. The four conditions varied with respect to the type and the amount of elaboration with which a list of 20 pictorial PAs were presented for learning.

Contrary to *a priori* predictions, middle-SES kindergarten and first-grade *S*s learned more efficiently in all conditions than did lower-SES children, whereas the performance of the two third-grade samples was equivalent in all conditions. Additional data from the present experiment along with relevant results of previous work were marshalled in support of interpreting the inferior performance of the young lower-SES children as an inability to benefit as much from repetition as did the middle-SES children.

◆ ◆ ◆

Racial Affect in Reading Comprehension

ROBERT L. AARON and WILLIAM F. WHITE, The University of Georgia

Kingston and White (1966) and White, Kingston, and Weaver (1967) investigated the relationship of judgments of the concept of self and concept of a protagonist (in a story) by means of the Osgood Semantic Differential (SD). Taylor (1957) introduced cloze reliability scores (CL) as indices of individual differences in comprehension and aptitude. The present study continued to examine the contribution of cognitive components (CL) and affective dimensions (SD) in readings about a Negro and a white male models.

Three groups of Negro fifth grade children, equated on intelligence and reading level, were differentially set to read about each of two stimulus figures. One group was shown a single picture of the Negro and White protagonists, while another group was given five pictures of game-like interaction of the two main characters with other members of a group. The control merely read the passage which contained no reference to the race characteristics of the protagonists. A factor analyses of SD responses rating the person-concepts of the Negro and white protagonists were obtained. Canonical correlation analysis was made from cloze test scores and factor scores on the SD.

1:00 to 1:55 P.M. **SESSION 12.8** **DIVISION A**

TEACHER BEHAVIOR AND ORGANIZATIONAL STRUCTURE

CHAIRMAN:

Robert W. Heller, State University of New York at Buffalo

PAPERS:

The Measurement of Teacher Dissatisfaction

RAY TOBIASON, Puyallup Public Schools Puyallup, Washington

The purpose of this study was to identify and measure dissatisfaction perceived by teachers. Hypotheses were tested to determine: (1) areas in which teachers perceive a lack of congruency between individual need and role expectation, and the magnitude of resulting dissatisfaction; (2) the relationship of magnitude of dissatisfaction in various organizational areas to the magnitude of dissatisfaction in areas outside the organization; (3) the relationship of dissatisfaction magnitude to membership in subgroups assigned on the basis of age, sex, perceived possibility of aspiration fulfillment in the present district, teaching level, and life-satisfaction sources; and (4) the feasibility of the Dissatisfaction Magnitude Scale (DIMS) as a diagnostic instrument for the measurement of dissatisfaction.

The Dissatisfaction Magnitude Scale (DIMS) was developed as a part of the study and administered to one-hundred classroom teachers selected at random and representing all grade levels in an 8,000 student suburban school district. The DIMS incorporated the bi-polar adjective design of the Semantic Differential utilized by Osgood, Suci, and Tannenbaum in the measurement of meaning. Dissatisfaction was measured on ten concepts present in the organizational setting (e.g., MY PRINCIPAL) and five concepts present outside the organization (e.g., WHERE I LIVE). Four factors, POTENCY, ACTIVITY, CONSISTENCY, and EVALUATION, identified through factor analysis of the twenty bi-polar scales, added to the DIMS's diagnostic strength.

Findings included: (1) individuals tended to maintain a consistent ratio of dissatisfaction between concepts present in and outside of the organization, but expressed a lesser magnitude of dissatisfaction on the latter; (2) dissatisfaction decreased with increasing age; (3)greater dissatisfaction was expressed by males, and by those with limited satisfactions coming from the organizational role; and (4) the DIMS shared equal validity with less diagnostic instruments commonly used in measuring dissatisfaction. It was concluded from these findings that the DIMS would be a device of high utility for administrators concerned with measuring staff dissatisfaction.

B. What is the effect of introducing a structured curriculum game to teach colors;

C. How do children in the preschool, who are not exposed to the curriculum game, progress in color learning;

D. Are there individual and group differences in learning to label colors, and are these differences related to the ages of the children.

An inventory which assessed the ability to match, receptively label, and expressively label colors, was administered to the experimental subjects and a comparison group. The experimental subjects then participated in a series of "fishing" games designed to teach the three levels of color knowledge.

The study suggests the following conclusions:

A. Few children from poverty homes can correctly label six basic colors when they enter preschool at the ages three to five;

B. These children do not rapidy learn to label colors in a preschool program that exposes them to many activities and experiences involving color, but does not specifically set objectives for color learning;

C. Introducing a structured game format to teach a series of precise objectives does produce comparatively rapid learning for many of the children;

D. There are large individual differences in the rate of learning shown by different children;

E. Some children do not learn from the kind of curriculum game employed here;

F. The amount and rate of progress made in color learning is not related to the age of the children in any clear-cut way.

◆ ◆ ◆

The Influence of Differential Community Perceptions on the Provision of Equal Educational Opportunities

JAMES G. ANDERSON, New Mexico State University, and DWIGHT SAFAR, Oregon Technical Institute

This study explores the perceptions of parents, teachers, school administrators, school board members, and the community at large as to the abilities of children from three ethnic groups: Anglo, Spanish-American, and Indian.

Major questions asked were: How does the dominant Anglo population view the abilities of Spanish-American and Indian children in comparison to their own? In turn, how do the Spanish-American and the Indian parent feel about the abilities of their own children in comparison to Anglo classmates? Is the failure of Spanish-American and Indian children in contrast to the relative success of Anglo children in school perceived as a result of differential treatment by teachers and school administrators? How much of this failure do the members of the three cultural groups and the educators themselves attribute to inadequacies in the school program?

In order to explore these questions, a study was conducted in two Southwestern communities. A two-stage cluster sample was drawn of families residing in each community. All available school board members and school administrators were interviewed. Also a sample of teachers from each grade level in each school was selected and interviewed.

The findings demonstrate a ubiquitous feeling that Spanish-American and Indian children are less capable of achieving desirable goals than are their Anglo contemporaries. This lack of achievement of the minority groups appears, in a large part, to be perceived as a lack of innate ability and support rather than as a fault of inadequate school programs. Moreover, this feeling of inferiority appears to be internalized by the minority groups themselves, thus creating an insidious negative climate for their children.

The result of this non-awareness of these differential community perceptions appears to be educational programs that are ill suited for Spanish-American and Indian children who enter school at a distinct educational disadvantage. Because of this gulf between the communities and their professional educators, the schools fail to assist the minority child in overcoming his educational handicaps and little is done to offer true equality of educational opportunity.

◆ ◆ ◆

Preliminary Findings from a Longitudinal Education Improvement Project Being Conducted for Instructionally Impoverished Pupils in Intact Schools in the Urban South

STANTON D. PLATTOR, Dillard University and Tulane University; and EMMA E. PLATTOR, Tulane University

A longitudinal study of teacher-learner interaction is being conducted, in two intact urban elementary schools, as part of the New Orleans Education Improvement Project (N.O.E.I.P.). The decision to conduct research in intact schools was made in order that conditions and activities could be replicated most effectively and efficiently, should the findings and conclusions of N.O.E.I.P. warrant such replication. The total population comprises an *N* of some 2200 pupils and some 90 teachers, administrators, and other personnel.

Changing Teacher Morale: An Experiment in Feedback of Identified Problems to Teachers and Principals

RALPH R. BENTLEY, Purdue University

The project reported here is primarily concerned with changing teacher morale. Given a certain level of teacher morale in a particular school situation, can the morale be improved by definite and deliberate procedures? More specifically, can feedback to the teachers and principal about problems and tensions as existing in their school situation be used to change morale? Is such feedback effective in stimulating individual and group efforts to alleviate tensions and overcome existing difficulties, thus bringing about an improvement in the level of morale?

Other purposes of the project include making morale comparisons between vocational and non-vocational teachers and determining whether a relationship exists between teacher morale and certain selected factors, e.g., sex, teaching experience, salary, teaching assignment.

The major questions to be answered in the study are the following: (1) Does feedback of teacher identified problems make a significant difference in changing teacher morale in particular school situations for (a) teachers generally, (b) vocational teachers, and (c) non-vocational teachers? (2) Do vocational teachers differ significantly from non-vocational teachers in the general level of morale and in terms of specific morale factors? (3) Is there a relationship between teacher morale and such factors as age, sex, teaching experience, level of education, salary, and major teaching assignment?

A two-year experimental study was conducted with the principals and teachers in 76 secondary schools in Indiana and Oregon. The total population consisted of 3070 teachers—223 vocational and 2847 non-vocational.

The Purdue Teacher Opinionaire was used for both pretests and post-test. The treatment for the experimental group consisted of feeding back to the teachers and their principal detailed information based on teacher responses to the Opinionaire made at the first pretest.

Analysis of covariance is being used to compare experimental and control groups.

◆ ◆ ◆

The Organizational Structure of Schools and Its Relationship to Teachers' Psychological, Sociological, and Educational Role Orientation

CHRISTOPHER FLIZAK, State University of New York, Fredonia

The organizational structure in which one works is supposed to have some effect upon the individual's mode of thinking, feeling, and behaving.

A theoretical model was developed defining three organizational structures of schools as *authoritarian, rationalistic,* and *humanistic.*

An empirical study was conducted to identify schools' organizational structure and then to ascertain any relationship existing between the organizational structure of a school and its teachers' mode of thinking, feeling, and behaving.

Thus, utilizing a set of criteria defining each model, four (4) out of 49 elementary schools were independently classified by three judges as representing each model. Teachers of these schools were given three tests. The Evaluation Modality Test aimed at ascertaining psychological factors on which their values were based, Teacher Practice Questionnaire aimed at their educational-role orientation, and Minnesota Teacher Attitude Inventory at their attitudes toward school situations.

A total of ten dependent variables from the three tests were examined in order to test for differences between the teachers belonging to different model schools. Seven out of ten dependent variables supported the hypothesis that the organizational structure and related dynamics of a school have significant relationship to certain social-psychological characteristics of its teachers.

Teachers of the AR schools attained significantly (.05) higher mean scores on the following measures: moralist and realist modes and on total intensity of the ethical valuation, and on advice information-giver, referrer, motivator and disciplinarian teacher-role measures. Teachers of the RH schools scored significantly higher on the individualist mode of ethical valuation, the MTAI, and on the counselor teacher-role measure.

The results of this study appear to support the idea that what happens in our schools may be determined by the very structure of these schools. Thus, in order to make any significant changes within the end-product of this process it may be necessary to examine organizational structure in which this process takes place.

◆ ◆ ◆

A Management Model for Research in Teacher Behavior

RUSSELL A. HILL, Temple University

This paper presents a systems management model for research in teacher behavior. The model's claim to utility lies in its ability to (1) systematically predict research problems; (2) specify resulting R&D tasks; and (3) allow the derivation of required finances and personnel needed to move forward. All of these tasks are organized into chronological stages of development wherein some tasks are prior and more significant than others.

The model is developed in four steps as follows:

1. Operational definition of the program objective
2. Definition of system units
3. Enumeration of the stages of system development
4. Construction of program phases as system units multiply

There are two dimensions of the model. The vertical dimension refers to unitary *system development*. The stages and sub-stages of system development are enumerated with the prior stages at the top and the latter stages of system development at the bottom as follows: theory, instrumentation, prescription, field implementation, and dissemination. The lateral dimension of the model refers to the program development as systems multiply. The units of this dimension are designated as *program phases*. Prior stages are at the left, and move to the latter stages at the right as follows: creative formulation, classification, comparison, synthesis, and stabilization.

Each vertical system stage and each lateral program phase is further analyzed into sub-objectives, required tasks, and needed resources. The result is a predictive description of the R&D program for the analysis of teacher behavior.

The utility of the model is tested (1) by using it to sort and to conceptualize current research activities in the field of teacher behavior, and (2) by projecting some hypothetically critical tasks for the future.

If the underlying assumptions and logico-deductive methods used are correct, the resulting model can provide a valid basis for the application of program planning and budget systems (PPBS) to teacher behavior. In effect, such a conceptual model provides a basis for decision making and planning for the manager of the R&D enterprise.

2:00 to 2:55 P.M. SESSION 13.2 DIVISION B

CURRICULA FOR DISADVANTAGED STUDENTS

CHAIRMAN:

Herbert M. Kliebard, University of Wisconsin

PAPERS:

The Effects of Individualized Instruction on Head Start Pupils' Achievement

KATHRYN B. DANIEL and MILLY COWLES, University of South Carolina

In the summer of 1967, 14 experienced teachers were enrolled in a special re-training program in early childhood education sponsored by the Southern Education Foundation. It was decided to select randomly 14 Head Start classes for the teacher-students to use as a laboratory in devising strategies for working with "disadvantaged children." (The criteria for inclusion in a Head Start center for the pupils were determined by OEO and are the same as are found nationally.)

The purpose of the investigation was to determine whether or not short periods of individualized training of five weeks (25 days) would be reflected in pupils' readiness as measured by the *Metropolitan Readiness Tests.*

Although the data were not significant at the accepted level (.05), they did show clearly that the pupils who received special instruction gained from the pre- to post- testing periods. Inspection of the data revealed that most pupils in control group either lost or remained at about the same level. Since the time span was so short, it is recommended that a similar experiment be conducted for various time periods, e.g., three, six, and nine-month periods.

The study was not designed to measure the growth of teacher-students; and it is further recommended that since, subjectively, skill appeared to increase substantially, further work be done to determine the value that may be derived in having teachers work specifically in situations in which they learn more about diagnosing pupils' specific instructional levels and devising strategies for appropriate treatments. A vast majority of the materials being marketed for use with young children are so "canned" that it becomes almost impossible for teachers to individualize instruction without special training.

◆ ◆ ◆

An Experimental Project to Teach Color Labeling to Disadvantaged Preschool Children

HARRIS W. STERN, Child Opportunity Program, Miami

The study sought to answer the following questions:

A. What is the level of color knowledge in disadvantaged children entering the preschool;

At the inception of N.O.E.I.P., the pupils displayed those learner behaviors characteristic of academically impoverished children in the urban South—including depressed learning potential, low readiness levels, inadequate reading skills, and poor in-school achievement. One on the basic premises established for N.O.E.I.P., therefore, was that a significant change in the pupils' academic potential (as measured by a standardized group test of intelligence) could be made as a result of appropriate modification of the learning environment in the two schools.

An examination was made of the pupils' entering behaviors in this area through the use of the Long Form of the California Test of Mental Maturity (CTMM). This test was selected because: (a) it is sufficiently long to provide a reliable measure of factors involved in academic potential; (b) it provides both Language and Non-Language I.Q. scores as well as Total scores; and (c) scores on this test are scaled to the Stanford-Binet, and therefore show similar relationships among CA, MA, and I.Q., as does the Stanford-Binet. (Limitations in terms of time and personnel made it impossible to consider annual pre- and post-testing with an individually administered instrument.)

The findings indicate that I.Q. scores as an index of academic potential are amenable to improvement when appropriate modifications of the learning environment are introduced. This improvement is accelerated when the general modifications are accompanied by the introduction of specific instructional interventions geared to the learning requirements of the pupils.

◆ ◆ ◆

A Study of the Effects of an Elementary School Enrichment Program on the School Achievement of Welfare Recipient Children

JOE L. FROST, The University of Texas at Austin

This study was primarily designed to determine the effects of an elementary school enrichment program on intelligence, personality, and achievement in language, reading, and arithmetic. The subjects were 574 first, second, and third grade children (222 welfare recipient) in three north central rural Arkansas schools. The experimental school (X) developed an "enriched" program of instruction over a ten-year period. The control schools (A and B) provided "typical" instruction. The California Test Series was administered for each academic area in question during September, 1966, and May, 1967.

The results revealed no pattern of superiority for either of the three schools, implying two alternatives: (1) Schools should produce *creative* compensatory programs based upon unique sets of assumptions, and/or (2) they should begin earlier. Welfare recipient children came from larger families than non-welfare children (seven to five, respectively). They performed significantly lower on tests of mental maturity, language, reading, and arithmetic with greatest differences in language and reading. Personality test performance was not significantly different. No pattern of achievement differences between sexes was revealed. Children enrolled in a special, small group intensive instructional program of reading and arithmetic in School A made gains superior to any other group studied. Comparison of this finding to results of large group enrichment suggests that small group intensive instruction is superior to large group "enrichment" or "regular programs" for compensatory programming. The findings suggest that comprehensive reexamination of the welfare condition should be undertaken for the possibility of isolation inherent characteristics of such status.

2:00 to 2:55 P.M. SESSION 13.3 DIVISION D

STATISTICAL TECHNIQUES

CHAIRMAN:

Louis L. McQuitty, University of Miami

PAPERS:

Some Varieties of the Linear Multiple Regression

C. E. ALEXAKOS, West Virginia University

The present study considers three basic models of multiple regression: (1) *Model A* (or complete model) where a matrix of correlation coefficients of order mxm is analyzed (1) to determine the beta coefficients and subsequently the value of the multiple R. In Model A no further analysis is undertaken; that is, all betas are retained. Model A has been described by Anderson (1958), Cornell (1956), and DuBois (1957) among many others; (2) *Model B* (or forward Model) where the analysis begins with a small number of predictor variables and continues by step by step addition of new predictors. Addition of each new predictor changes the values of the regression coefficients. In this model significant predictors are retained while non-significant predictors may be dropped and new ones added until the best regression equation is obtained. Model B is discussed by Fraser (1958) and Friedman and Foote (1957); (3) *Model C* (or backward model) where the analysis commences with a complete matrix of correlations, but does not terminate at the first step. In subsequent steps, non-significant predictors are eliminated or suppressed one by one. The classical Model C is described by DuBois (1957), Johnson (1949), and Walker and Lev (1953).

The classical Model C (designated as C-1 in this paper) follows the opposite procedure of Model B. In B, predictors are added one by one; in C-1 predictors are eliminated one by one. The three models A, B. and C-1 can be considered as orthogonal models.

In the present study a distinction is made between orthogonal models and oblique models. The backward solution seemed more suited to this distinction than the forward model. The backward C-1 model gives identical regression coefficients at corresponding steps with the forward model. If, however, instead of simply eliminating a non-significant predictor from the matrix of intercorrelations, we partial it out (i.e., actually compute the partial correlation coefficients) in the matrix of predictors, we introduced a new model, termed Model C-2 in this study.

Further, if we compute the partial correlation coefficients, partialling out a non-significant predictor in the original matrix R including the criterion variable, we introduce a new Model C-3.

Model C-2 may be found useful in studies where the predictors are correlated among themselves but are independent of the criterion. Model C-3 may be used when both predictor and criterion variables are dependent; that is, in cases where the investigator cannot experimentally control the criterion variable.

Application of the five models distinguished here to the same data produced three different sets of beta coefficients. Models A and C-3 gave identical values of beta. Also, Models B and C-1 produced a second set of b_i, but different from the set given by Models A and C-3. Model C-2 produced a third set of coefficients entirely different from the other two.

Further research may be necessary on the oblique Models C-2 and C-3. It was the intent of the present paper to generate discussion on the last two models and interest in exploring their usefulness in educational research.

◆ ◆ ◆

Canonical Regression: An Extension of the Procedure of Canonical Correlation

C. H. BRADFORD, Washington University

The paper outlines some theoretical and empirical explorations of an extension of the procedures of canonical correlation. These extensions of canonical correlation into canonical regression equations yield results that are comparable to and as intuitively direct as the interpretations of multiple regression equations. The similarities and differences in the results of these two procedures are illustrated using data relating teacher personality need profiles to teacher attitudes toward pupils. The data used was gathered over two summers using the *Edwards Personality Preference Schedule*, the *Minnesota Teacher Attitude Inventory* and the Wehling-Charters' "Teacher Conceptions of the Educative Process."

The results of canonical correlation procedures where the maximum number of canonical variates have been extracted can be expressed in matrix form as follows: (1) $By = Ax$. Where B is $p \times p$; A is $p \times q$; y is $p \times 1$; x is $q \times 1$. In this form the corresponding rows of A and B are the coefficients of the associated canonical variates. If p is less than or equal to q then the matrix B is non-singular, hence (2) $y = Cx$. Where $C = B^{-1}A$. The y in equation (2) can be viewed as the values predicted by the set of canonical equations given a particular set of x values. C is then the matrix of canonical regression coefficients, a matrix counterpart to the vector of regression weights in the usual multiple regression. If y is interpreted as the set of dependent variables and x as the independent set, the equation (2) yields the best joint prediction of the dependent variables y from the set of x values, subject to the usual conditions of canonical correlation.

The matrix C has a further, more directly interpretable meaning. If A and B in equation (1) are expressed in standardized form then column j of the matrix C is the predicted pattern of the dependent variables when the independent variable has a standard score value of +1.0 and the remaining q — 1 independent variables have standard score values equal to their mean, zero. Further row i of matrix C is the canonical counterpart of the beta weights of the multiple regression of dependent variable i on the independent set.

Since their introduction by Hotelling, the procedures for canonical correlation have been seen by many as a more realistic model of the conditions of multiple determination than the multiple regression procedures which ignore any multiple dependencies in the dependent variables. However, a major drawback to the use of canonical correlation has been the difficulty in interpreting the canonical variates in any intuitively meaningful terms. The above mathematical extensions of the canonical equations, which to my knowledge have not been used before, make it possible to interpret canonical correlations in terms that are as direct and also are comparable to the interpretations in multiple regression.

◆ ◆ ◆

Converting an Inconsistent Matrix into a Consistent Non-singular Matrix

MARION F. SHAYCOFT, American Institutes for Research

When correlation matrices are used for multiple regression, factor analysis, canonical correlations, or any other form of multivariate analysis, it is of course important to have the correlation matrix internally consistent—in other words, Gramian. Unfortunately this isn't always feasible. Sometimes there is an extensive problem of missing data, making it impractical to base all correlation coefficients on exactly the same cases. This may produce a non-Gramian matrix. Correcting correlation coefficients for attenuation may also have this effect, if reliability coefficients that are underestimates are used. A non-Gramian matrix results in impossible values, such as multiple or partial coefficients above 1 or imaginary.

To cope with these problems, the author developed a statistical procedure for making minimal modifications in any matrix that proves to be inconsistent, to convert it to a consistent matrix. Operational use of this procedure has shown it to work very well. The chief drawback was that the revised matrix was necessarily singular, a characteristic that resulted in computational inconvenience and generally necessitated the elimination of one or more variables before multivariate analysis could be undertaken on the corrected matrix.

The drawback noted above has now been eliminated by a simple modification in the correction procedure. This modification makes it far more likely that the corrected matrix will be non-singular.

Results are presented for a matrix that was made inconsistent by deliberately using gross underestimates of reliability for some variables, in correcting for attenuation. The revised procedure will be shown to produce a matrix that is both nonsingular and very close to the one that use of more accurate reliability estimates would have produced.

◆ ◆ ◆

Blown-up and Regression-derived r's

JULIAN C. STANLEY, Johns Hopkins University

Glass (1966) derived a formula for the correlation between a set of consecutive, untied ranks 1, 2, . . ., n and a two-point distribution under the assumption that a rectangular distribution of the ranks underlies the two-point distribution. In this paper it is shown that this coefficient is equivalent to the computed Pearson product-moment r between the ranks and the two-point distribution, divided by the maximum value that such an r can attain, and also that the usual biserial r can be computed as point-biserial r divided by the maximum point-biserial r that can occur when the scores of the multicategoried distribution are normally distributed. It is pointed out that not all commonly used correlational coefficients have this property. For example, tetrachoric r is not a divisor-inflated four-fold-table product-moment r, nor is the obtained contingency coefficient divided by the maximum possible contingency coefficient for a contingency table of a certain number of rows and columns derivable by Pearsonian linear regression-line methods.

2:00 to 2:55 P.M. SESSION 13.4 DIVISION D

MEASUREMENT OF COMPLEX INTELLECTUAL PROCESSES

CHAIRMAN:

Dorothy M. Knoell, American Association of Junior Colleges

PAPERS:

Relationships Among Selected Measures of Creativity

W. SCOTT BOWER, The Ohio State University

Creativity research has received much recent criticism. Probably the most cogent attacks have been those directed at the inability of researchers to develop measures of the proposed creativity dimension which relate highly to each other. It has been demonstrated that so-called measures of creativity have, in some cases, actually been more highly correlated with commonly used measures of intelligence than with each other.

The present study was undertaken in an effort to determine whether an interpretable pattern might be found among a number of frequently used group measures of creativity. Nine instruments were chosen which could be variously scored so as to yield eighteen scores. Represented among the instruments are both verbal and non-verbal tasks and both cognitive and affective measures. Some of the measures are based upon relatively well developed theoretical positions, while others have apparently been used on the basis of surface considerations. Some have been more thoroughly empirically validated than others. These tests (plus a verbal and a non-verbal measure of intelligence inserted for purposes of comparison) were administered to a group of sixty "Upward Bound" students. Some were administered during the summer of 1966; others during the summer of 1967. Scores were intercorrelated and the matrix was factored by the principal axes method and rotated to simple structure by means of Kaiser's Varimax technique.

Eight factors accounted for all of the common variance. One of these was clearly an intelligence factor. The other seven were discussed in terms of the pattern of relationships represented by them. It was concluded that, while it is difficult to defend a unidimensional conception of creativity on the basis of present evidence, perhaps the chief problem in creativity research is lack of concensus as to appropriate measures. A more theoretically oriented approach to the development of creativity measures was recommended.

◆ ◆ ◆

Diagramming as a Means of Representing Structure of Knowledge and Testing for Cognitive Structure

JEAN DYER, University of Alberta

Structure of knowledge is defined as the organization of a given area of knowledge, where organization refers to the relationships between elements (concepts, principles, events, theories, substructures, etc.) within that area. Most relationships can be classified as follows: descriptive, inclusion-exclusion, causal, temporal, logical, quantitative, and functional.

Diagrams are proposed as a means of representing structure of knowledge. There are at least three basic diagrams: Venn diagrams (represent inclusion-exclusion relationships); tables or matrices (represent descriptive and causal relationships); and graphs, i.e., graph theory, points and lines (represent causal, temporal, logical, quantitative, and functional relationships). Therefore, diagramming a given area requires systematic definition of the basic concepts and existing relationships between them, followed by selection of an appropriate diagram form.

This diagram of the structure of an area also serves as a blueprint for constructing tests, which measure cognitive structure relationships and determine how well the structure of the material has been transmitted to the student's cognitive structure. If both cognitive elements and relationships were tested an essay test would probably be required. Instead, only relationships are systematically examined within an objective format where the elements are given. Each item or group of items is analyzed for the number of relationships tested and scored accordingly. Items are often scored on the basis of patterns of responses, different patterns reflecting different cognitive structures, rather than scoring each item independently. If test items are constructed to yield enough information, it is possible to "diagram" an individual's cognitive structure.

Sequential or learning dependencies within the material (Guttman scale patterns) can be pinpointed from the structure of knowledge diagram. Transfer and inference items are easily generated. The diagram procedure also has implications for individualized instruction because emphasis is on mastery of subject matter, not relative achievement among students as is the case with most achievement tests.

◆ ◆ ◆

Association Value and Subjective Ratings of Interest in Visual Complexity

D. R. EVANS, and H. I. DAY, Ontario Institute for Studies in Education

In several studies Day has used random polygons to measure interest in visual complexity (viz. Day, 1967). As with nonsense syllables random figures have been shown to have association values (Vanderplas and Garvin, 1959). This study was designed to investigate the relationship between the association value and subjective ratings of interest in a set of random figures of low, medium, and high levels of complexity and the effects of familiarity on these response measures.

Sixty-four subjects divided into two equal-sized groups were shown 75 random figures; 25 at each of three levels of complexity (10, 28, and 80 sides). Slides of the figures were projected in random order by a Kodak "Carousel" projector on a 48″ x 48″ screen. During the association task each slide was exposed for 15 seconds and subjects were asked to record all their associations to the slide. In the interestingness task the slides were projected for 15 seconds and the subjects were asked to rate the interest level of each figure on a seven point scale. Order of tasks was reversed for the two groups to examine the novelty/familiarity effect. Associations and interest ratings were summed across subjects and a Pearson r was calculated at each of the three levels of complexity. For the low complex material (10 sides) r was .35, for the medium complex material (28 sides) r was .71, and for the high complex material (80 sides) r was .52. On both tasks subjects tended to obtain higher scores when the material was novel than when it was familiar. This finding was most apparent for the high complex material.

The results of this study were interpreted in terms of Berlyne's (1963) theory of subjective responses to complex situations.

◆ ◆ ◆

A Correlational and Factor Analytic Study of Three Critical Thinking Tests, One English Test, and One Logical Reasoning Test

JOHN FOLLMAN, University of South Florida

A review was conducted of the theoretical and empirical critical thinking, reasoning, language and thinking literature. It appears from this review that critical thinking is an important educational objective, that it has received substantial empirical and theoretical attention, and that the nature of critical thinking is amorphous, confusing, complex, and unclear. It also appears that the relationships between and among critical thinking, reasoning, and language, are also unclear.

The objective of this investigation was to attempt through correlational and factor analysis to obtain better understanding of the nature of critical thinking and the inter-relationships between critical thinking, reasoning, and language.

Two series of analyses were conducted. One series consisted of four factor analyses, one of a matrix of all twenty-four subtests from the *Watson-Glaser Critical Thinking Appraisal: Form ZM: The Cornell Critical Thinking Test: Form Z;* American Council on Education *A Test of Critical Thinking: Form G; Logical Reasoning;* and *The Cooperative English: Form 1B.* Also analyzed were a 23 × 23 matrix of all subtests except *Logical Reasoning,* a 22 × 22 matrix of all subtests except the two *COOP* subtests and a 21 × 21 matrix of all subtests except *Logical Reasoning* and the two *COOP* subtests. The Pearson Product Moment intercorrelations matrices were analyzed by Principal Components factor analysis and Kaiser Normalized Varimax Rotation. Ones were placed in the diagonals and rotation was continued until seven or more factors were obtained.

The other analysis was a factor analysis of a 5 × 5 matrix of Pearson Product Moment intercorrelations of the five total test scores.

In the subtests' analyses while no large general factors were found a considerable number of distinct specific factors were identified.

The analysis of the 5 ×5 total test intercorrelations indicated consistency between the critical thinking tests, the critical thinking tests and the English test, and a separate distinct *Logical Reasoning* test grouping.

2:00 to 2:55 P.M. **SESSION 13.5** **DIVISION C**

VERBAL BEHAVIOR IV

CHAIRMAN:

Walter H. MacGinitie, Columbia University

PAPERS:

Dimensions in the Rating of Compositions

J. M. SCHUERGER, Western Reserve University and R. G. FRANKIEWICZ, Kent State University

The purpose of this study was to determine the extent to which experienced English teachers could rate short compositions on hypothetically independent dimensions. The ratees were 177 ninth-grade boys in three different schools. Homogeneity was assumed from similarity of socioeconomic status and absence of significant inter-school differences on 11 of 15 aptitude tests.

The students wrote twenty-minute compositions on two topics, one designed to elicit expository prose, the other imaginative prose. Each composition was rated by three readers (scale 1-7) on five hypothesized dimensions: idea, the content of the composition; form, the organization of the ideas; flavor, personality of the writer as expressed in the composition; mechanics, aspects of composition such as grammar and spelling; wording, use of words. During training, readers were instructed to rate all compositions for one dimension before proceeding to the next dimension.

For each of the five dimensions on each topic sums of the three scale values assigned by the raters were intercorrelated; the resulting 10 × 10 matrix was factored by an iterative procedure initialized with unities in the principal diagonal. The sequence and size of the eigenvalues indicated that interrelationships among the sums could be explained in terms of three factors. The three factors were rotated to the varimax criterion.

The results suggest that, with the exception of mechanics, dimensions of composition skill should be thought of in terms of the kind of topic used as stimulus rather than as skills of composition across topics.

◆ ◆ ◆

A Method for the Analysis of Language Habits in Science

RAYMOND E. SANDBORGH and PAUL E. JOHNSON, University of Minnesota

The description of intra-verbal relations among concept words in science is one means of making predictions about the role of language habits in the acquisition and retention of such concepts. Existing methods based upon commonality among free-association responses allow for the prediction of co-occurrence relations among stimulus words (Deese, 1965; Garskoff and Houston, 1963). However, since many patterns of usage in the language of science involve more than two concept words, a more general means of description is needed. This paper presents one such description based upon a set theory formulation of word association data which reduces to the coefficient of relation for two stimulus words as a special case. Data are presented to illustrate the description for a well-defined set of stimulus words.

◆ ◆ ◆

Effects of Written Instructional Material on the Statistical Structure of Test Essays

ERNST Z. ROTHKOPF, Bell Telephone Laboratories, RONALD THURNER, University of Minnesota, J. DOUGLAS CARROLL, Bell Telephone Laboratories

Essay examinations have some of the properties of the Free Recall procedure that has found wide use in laboratory studies of verbal learning. Structural information has proved to be one of the most interesting characteristics of free recall protocols. The purpose of the present experiment was to explore some techniques of analysis of the structure of essays. It was hoped to find new means for measuring outcomes of instruction and of other learning experiences.

High school students wrote short essays on each of 19 concepts from Newtonian mechanics both before and after reading a relevant 20,000-word physics passage. A control group read a passage of equivalent length from a biology text.

The frequency of key concept words as well as co-occurrence of pairs of words in the same context was determined by a computer analysis of the essays. A similar analysis was performed on the instructional text. Psychological proximity in a two-dimensional space, was computed by means of the parametric mapping technique devised by Carroll. This technique results in the determination of proximities between two words as an increasing function of the frequency of co-occurrence of the two words within the same context, i.e., the same sentence.

The structure of essays written after reading the physics text, as measured by proximities among the 19 key concept words, tend to resemble the text structure more than those of essays that had been written without seeing the instructional text. Proximities based on raw co-occurrence frequencies tend to be somewhat more sensitive measures of instructional effects than the scaled proximity data, but the scaled data is a more convenient and succinct characterization of instructional outcomes.

Statistical characteristics of word distributions in essays appear to have some potential usefulness as indicators of the effectiveness of instructional documents. Among the advantages of such measures is that many of them can be carried out automatically by computers.

◆ ◆ ◆

Some Cognitive and Affective Variables as Correlates of Connotative Meaning

WENDELL WEAVER, WILLIAM F. WHITE, ALBERT J. KINGSTON, The University of Georgia

Although there are referential aspects to "connotative meaning", the construct has generally been measured by Osgood's Semantic Differential which is an instrument focusing on the affective domain. This study involved the use of the semantic differential as a criterion variable, and examined the amount of variance in the prediction of the criterion that was attributable to separate cognitive and affective variables.

Cognitive characteristics were indexed by the cloze procedure—a language deletion technique in which every nth word of a natural language passage is deleted. The Ss task was to supply the exact word removed from the passage. Pre-cloze (i.e. before the Ss read the passage intact) tests have been shown to correlate highly with measures of intelligence (in the 70's) while post-cloze tests have been correlated highly with paragraph comprehension tests and multiple-choice achievement tests (in the 80's and 90's).

Additional affect aspects were indexed by the second-order factor scores of the 16-PF. Cattell's labels were used for identifying these four dimensions—anxiety, extroversion-introversion, tough poise, and independence.

Ss were 121 female and 41 male upper classmen at the University of Georgia enrolled in educational psychology courses. The tasks were given in the order: SD-self, pre-cloze, 16PF, reading passage, SD-protagonist, and post-cloze. Three separate periods were used for the testing.

Results were analyzed separately for males and females, SD-self, and SD-protagonist. It appeared that for female subjects only, both semantic differential concept measures, the 16PF second-order factors, and the cloze measures were not all related to each other in the same construct system. Both the 16-PF second-order factors and the cloze tasks were related to the semantic differential concept measured. Relationships are complex and are discussed at length. The results of this study are to be used to direct further research in which variables are to be manipulated. The directions of this research were discussed.

2:00 to 2:55 P.M. **SESSION 13.6** **DIVISION C**

LEARNING AND INDIVIDUAL DIFFERENCES II

CHAIRMAN:

Henry T. Lippert, Florida State University

PAPERS:

Classification Skills in First Grade Children: The Effects of Different Instructional Methods

MASAKO N. TANAKA, Educational Testing Service

The development of appropriate instructional methods involves a consideration of the information processing systems available to the child at the time of instruction. This study investigated the effects of two instructional strategies (object-manipulation and picture-verbalization) on the development of classification skills in first grade children. It was hypothesized that each of the two strategies would be effective with different initial levels of classificatory ability, i.e., the "verbalization" method would increase the scores at the upper part of the range more than the scores in the lower part of the distribution, whereas the converse would occur for the "manipulation" treatment.

The subjects included three first-grade classes (two experimental, one control) in a predominantly Negro, lower-middle class, urban school. The investigation covered a period of two weeks with eight instructional sessions and two testing sessions, each requiring approximately thirty minutes. The investigator administered all the tests and conducted the instructional sessions for both treatment groups.

The results indicated that the treatment groups made greater gains in mean scores than the control group. There was a significant difference in performance ($p < .025$) with most of the difference attributable to the superiority of the "manipulation" treatment over the control group.

The effects of the two treatments on differing initial levels of classificatory ability were investigated by a regression analysis of the posttest on pretest scores. There was a significant difference for the "verbalization" treatment ($p < .001$) in the hypothesized direction, i.e., the "verbal" approach increased the scores at the upper part of the distribution more than those toward the lower end of the score range. The intersection of the regression lines for the two treatments near the mean pretest and posttest scores suggests the usefulness of developing instructional strategies which are appropriate to the needs of different groups of children.

The Effect of Student Characteristic-Teaching Method Interactions on Learning to Think Critically

JAMES P. SHAVER, Utah State University

Research in teaching methods has not had a particularly fruitful history. Undoubtedly, this is due in some measure to factors such as inadequate criterion measures (McKeachie, 1962) and the failure to confirm that differences in teaching method did actually occur (Wispe, 1953). But to a large degree, the inconclusive and inconsistent findings must be seen as the outcome of research strategies that have failed to take into account that averaging scores and comparing means obscures the differing affects the same method may have on different students.

A project to develop materials for teaching junior high school students to analyze public controversy more adequately, has been reported previously (Oliver and Shaver, 1966). The evaluation was designed to make the usual experimental control group comparisons, and also to test the relative effectiveness of two different types of discussions used with the experimental students. One, the "socratic," was to be a probing dialogue in which the teacher would force individual students to take a position on a public issue and deal with inconsistencies in belief and attitude while defending their stand. The "recitation" discussion was aimed at a general discussion of the same issues, but without forcing students to take personal stands and defend them. The success of the experimental teachers in using the two discussion styles was confirmed using systematic observations (Shaver, 1964).

On a number of measures of learning of analytic concepts, the gains of the experimental students significantly exceeded those of students in the control schools, but in only one instance was there a significant difference between the two discussion styles.

In the traditional methods research study, this would have been the terminal finding. However, a number of personality measures had been administered to the experimental students ($N = 125$) in order to investigate student characteristic-teaching method interactions. Distributions of scores on the personality measures were cut into thirds, students categorized by discussion style and their place in the distribution, and two-way analyses of covariance (adjusting for pretest and mental ability scores) carried out. Twelve significant interactions were found. Tables presenting these will be included in the paper along with descriptions of the variables involved.

Patterns that emerged will be discussed, including two findings with important implications: (1) the number of significant interactions increase as the criterion measures became less direct (and probably more valid) in their measurement of learning; (2) frequently, students in the low and high thirds of the distribution on a personality variable reacted similarly to a discussion type, with students in the middle of the distribution reacting differently as indicated by test performance.

In conclusion, the amount of information acquired by proceeding to an analysis of student characteristic-situation interactions will be stressed. However, the difficulty of investigating these interactions given the complexity of the phenomena being studied and the need for programatic research to build cumulative, replicated data will be emphasized. (Reference will be made to a study to be completed in 1969 which is aimed at extending the findings discussed in the paper.)

◆ ◆ ◆

Individual Differences and Instructional Film Repetitions

RICHARD E. SNOW, Stanford Center for Research and Development in Teaching; **WARREN F. SEIBERT**, Purdue University; and **ROBERT W. HECKMAN**, Ford Motor Company

In experiments using laboratory learning tasks, learning can typically be defined in terms of improvement with practice; individual differences in learning can be represented by trial scores, and can then be related to other learning performances or other personal characteristics of learners. With this approach, changes in the structure of performance as a function of practice can be investigated, and reference abilities can be used to trace such changes during acquisition. In research on formal instruction, however, trial data have not normally been obtainable. Pre vs. posttest comparisons or gains scores derived from these tests have had to suffice. Thus, the presence or absence of some degree of learning has been studied, but there have been few attempts to investigate individual differences during intermediate stages of meaningful learning. The present study was designed to examine the characteristics and correlates of practice scores obtained from repetitions of instructional films and associated achievement tests.

A battery of reference tests was given to 185 undergraduates who were then randomly divided into two experimental and two control groups. Experimental Ss received instruction via two 30-minute sound films, with order of presentation counterbalanced between the two groups. Five film repetitions, one every 24 hours, interspersed between six repetitions of a criterion achievement test, were administered for each film. Control groups received repetitions of the achievement measures only, with order of presentation similarly counterbalanced. Test items were presented on slides to control exposure time. Analyses of variance testing experimental, order, and practice effects, as well as correlational and factor analytic methods, were used.

Intrafilm trial intercorrelations displayed simplex patterns for both experimental and control groups, but only the experimental matrix contained significant interfilm correlations. Significant correlations with film trial scores were obtained for tests representing Guilford's semantic cognition, memory, and convergent thinking categories, and for some newly constructed audio-tape and film tests of memory skills. Positive correlations were also obtained for reservedness, self-sufficiency, and several other scales from Cattell's 16 PF inventory. Results were fairly similar across films and trials within films, though some systematic variations were noted. Marked differences were observed between experimental and control groups.

◆ ◆ ◆

Fourth Grade Achievement as Related to Creativity, Intelligence and Teaching Style

GLEN ROBBINS THOMPSON, Skokie School District 68 and
NORMAN D. BOWERS, Northwestern University

This study was designed to investigate the degree to which such factors as student creativity and intelligence interact with teaching style to influence level of academic achievement.

Fourth grade children were tested in the Fall with the Lorge-Thorndike Intelligence Test and the Minnesota Test of Creativity. Then teams of observers rated each teacher's teaching of a 30 minute standard curricular unit in social studies, using the OSCARs 4v and 2e. Data pertaining to teacher supportiveness, verbosity, and convergence-divergence were thus obtained. At the end of the school year, all students were given the Stanford Achievement Test. Stanford Social Studies and Word meaning scores were related to I.Q., creativity and teaching style; and IQ and creativity test scores were used in combination to predict academic achievement.

Lorge-Thorndike IQ was significantly related to both Social Studies and Word Meaning scores on the Stanford. In addition, Stanford Word Meaning scores were found to be significantly related to certain other factors in the experiment. A two-way interaction was found between teacher verbal output and IQ. Students with "low" IQ showed the greatest achievement with either high or low verbal teachers, and the least with teachers who were medium in this respect. "High" IQ students, on the other hand, showed the greater achievement with medium, but especially high verbal teachers. A three-way interaction was found between teacher convergence-divergence, IQ, and creativity. "Low" IQ, "high" creative children scored highest on Word Meaning, if they had teachers that were neutral along a convergence-diverence continuum. This aspect of teaching style was not a relevant factor with children that were "low" IQ, "low" creative. In contrast, "high" IQ, "low" creative students scored highest with teachers neutral in this regard, while "high" IQ, "high" creative scored highest with convergent teachers.

Creativity test scores did not add significantly to the precision of prediction when used in a multiple regression equation with Lorge-Thorndike

IQ. This was true for both Stanford Word Meaning and Social Studies scores.

The results of the experiment are discussed in relation to previous research findings regarding teacher characteristics, intelligence, and creativity.

9:00 to 9:55 A.M. SESSION 14.1 DIVISION C

JOINT SESSION WITH NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION TECHNIQUES FOR INTERPRETATION OF RESULTS

CHAIRMAN:

Robert Lankton, Detroit Public Schools

PAPERS:

Item Analysis Procedures Used in the Development of Tests for the National Longitudinal Study of Mathematical Abilities

L. RAY CARRY, The University of Texas and JAMES W. WILSON, School Mathematics Study Group

At The School Mathematics Study Group (SMSG), we have selected, borrowed, and developed procedures for the evaluation of the statistical nature of a scale composed of items. This paper describes and illustrates these procedures.

The National Longitudinal Study of Mathematical Abilities (NLSMA) required development of a wide variety of tests. During the period since 1961, a continuing effort of SMSG has been the development of tests for NLSMA and other SMSG research projects. The primary importance of this work is the identification of appropriate behaviors for testing, and the construction of valid measures of these behaviors. A secondary consideration, but one necessary to the successful completion of this effort, is the examination of the psychometric properties of these measures, and the use of this statistical information for developing the final version of the scales.

This paper describes selection of appropriate statistics, the computations and displays, and the development of decision rules for using this statistical information in our test construction. Examples from the SMSG-NLSMA test construction activity will be used to illustrate various points in the paper.

◆ ◆ ◆

Test Score Profiles 1984

HARRY J. CLAWAR, Hunter College and Educational Records Bureau

Most test publishers and users are aware that all test scores contain certain amounts of error. It seems reasonable that scores should be reported in a manner that allows for the unreliability of the measuring instrument.

Two well-known approaches to the problem are represented by the STEP profile and the DAT pupil profile chart. The use of the aforementioned profiles, however, tends to be somewhat cumbersome for the amount of information obtained. For the DAT profile, one must measure differences between the height of subtest bars, and find them to be one inch or greater in order to consider the subtest differences as real. For the STEP profile, one must visually compare test score bands to judge the extent of overlapping (no overlap equals an important difference). At the end of both procedures we have information about differences among a pupil's set of scores.

A new profile system (ERB pupil profile chart), yields not only differences among all possible pairs of subtests, but also pupil vs. class median and pupil vs. norm-population-median differences. All these comparsions are made by computer routine (Clawer, 1967), using appropriate standard error formulae (Davis, 1959), and presented in a highly attractive and simplified format. Differences at the 15 per cent level of significance are indicated by the presence of symbols at designated profile locations.

Also generated at the end of each group (class, grade, etc.) of individual profiles is a summary profile. This profile indicates significant differences between group medians and norm medians as well as for all possible group median comparisons.

◆ ◆ ◆

An Exploration of the Utility of Self-Reported Test Scores

ROBERT J. PANOS, American Council on Education

In many research applications, it is desirable to include information concerning individual test scores among the variables to be examined. In most instances this may involve a simple clerical or computer look-up of available test scores followed by a file update. However, the task becomes relatively expensive and clerically cumbersome when the file is large. Furthermore, when test score data are not readily available, constraints such as obtaining clearance to access the data and increased costs, often force the researcher to forego utilizing difficult to obtain test scores, or otherwise to alter his research design.

Although researchers have been willing to accept and to utilize student self-reports regarding biographic and demographic data and, in many cases, student self-reports concerning objective achievements and attained grades, little attention has been devoted to the possible utility of self-reported test scores in educational research. The main objective of this study was to explore the usefulness of self-reported test scores compared with scores of record.

Self-reported and actual SAT test scores were available for a sample of 2,318 freshmen, representing the entering freshmen classes at eight women's colleges. Self-reported data were obtained from the students at the time of fall registration in 1966 as part of the American Council on Education's Cooperative Institutional Research Program; grades of record were obtained in early Spring, 1967. The sample was reduced to 1,212 students, since only those students for whom both self-reported SAT-verbal and SAT-math scores were available are included in the analysis reported here.

Intercorrelations, computed among self-reported and actual scores, showed a remarkable degree of overlap. The correlation between actual and self-reported SAT-verbal and SAT-math scores was .92 and .91 respectively over the entire sample. Correlations ranged from a low of .82 to a high of .94 in the eight colleges. In the paper, methodological problems are discussed, and implications concerning the possible utility of self-reported test scores in educational research are examined.

9:00 to 9:55 A.M. **SESSION 14.5** **DIVISION C**

CONCEPT LEARNING III

CHAIRMAN:

Marvin D. Glock, Cornell University

PAPERS:

Extra-Scope Transfer in Learning Mathematical Rules

JOSEPH M. SCANDURA and JOHN H. DURNIN*, University of Pennsylvania

A restricted rule-statement was defined as one obtained by replacing the response determining variables in a general rule-statement with the specific values of a particular instance. It was hypothesized that a restricted rule-statement might well provide a basis for generalization to all problems within the scope of the corresponding *un*restricted rule. It was further hypothesized that if performance on one extra-scope problem indicates that a restricted rule has been generalized, then transfer to additional problems within the scope of the more general rule should also be expected.

66 high school *S*s were taught a restricted statement (S', SG', or G') of one of three rule-statements of varying generality, S (= S') < SG (SG') < G (G'). 22 *S*s served as a control (C). All *S*s were tested on six problems, the first two within the scope of the most specific rule (S), the second two within the scope of only the more general rules (SG and G), and the last two only within the scope of rule G. Statement S', SG' ,and G', were directly applicable only to the first two problems. Groups SG' and G' evidenced extra-scope transfer. Groups S' and C did not. In addition, performance on the second problem of each pair was contingent on performance on the corresponding first problems, indicating that "what is learned" may be determined by performance on single test items and used to predict performance on additional similar-scope problems. Suggestions were made for future research.

◆ ◆ ◆

Relatedness and Dominance of Negative Instances in Concept Attainment

REED G. WILLIAMS, Indiana University (William W. Lynch, Indiana University)

The purpose of the present study was to investigate the joint effects of relatedness and dominance in negative instances on concept attainment. (Relatedness refers to the closeness of the relationship *between* the negative instances. Dominance pertains to the *strength* of the individual negative instance to elicit an implicit response from the subject). This study is an elaboration of an earlier study by Kendler and Karasik (*J. Exp. Psychol.*, 278-283, 55, 1958) which did not include the full range of combinations of relatedness and dominance, and which utilized an analysis not permitting the isolation of interaction effects.

Each of 35 subjects attempted six concept attainment tasks, permitting exposure to all possible negative instance conditions (high, middle, low dominant related, and high, middle, low dominant unrelated). Materials used were words taken from the Underwood-Richardson norms. The concepts were represented by four related words (positive instances) in the mid-dominant range. Sets of four negative instances were randomly paired with the six sets of positive instances. Two response measures, time to concept attainment, and errors to criterion, were used. Analysis of variance indicated significant differences at the .01 level for dominance and for interaction on both measures, while there were no significant differences for relatedness.

Kendler points out that differentiation of negative from positive instances is the prime task in concept attainment. He suggests that high dominant negative instances facilitate differentiation by *strongly* eliciting implicit responses unrelated to the concept. Relatedness facilitates concept attainment by collectively emphasizing the differentiation of the negative instances from the positive.

This study tends to support Kendler's findings but suggests that his interpretation needs revision in order to account adequately for the present results.

◆ ◆ ◆

Assessing Subject Matter Concepts: Relationships Among Item Forms and the Effect of Sequence

STANLEY L. DENO, University of Delaware

Generalizations regarding human conceptual behavior are frequently made from experiments in which concepts are measured through the operations of classification and definition. Yet few, if any, attempts have been made to establish a continuity between experimental studies of concept learning where the criterion task is classification, and the learning of subject matter concepts where the criterion task is rarely classification. The present study was designed to yield data bearing on several issues. First, are classifying, defining, and generating novel examples, equivalent tasks in the sense that each is a behavioral concomitant of the same cognitive process? Second, is there a prerequisite or contingent relationship among these three commonly used indices of concept attainment? Third, can these widely used experimental tasks serve as paradigms for developing item forms (types) when testing for the attainment of complex subject matter concepts? Fourth, does the sequence in which these item forms are encountered during testing influence achievement?

Students in undergraduate educational psychology class served as subjects for the research. As part of a regular class assignment the subjects were given a list of concepts commonly taught in that section of the course dealing with the psychology of learning, and were informed that in one week they would be tested to determine whether or not they had learned those concepts. A test was constructed so that for any concept to be learned the subject would classify, define, and generate novel examples for that concept. After one week all subjects were tested and the results were analyzed to determine the relationships among performances on the various tasks and the effect of task sequence during testing.

The results indicated the following: 1) Differential relationships among item forms suggest that quite distinct cognitive processes are involved; 2) defining and generating novel examples are closely related cognitive activities, but classifying is not; 3) experience on the classification task facilitates performance on the defining and novel example generating tasks; 4) subjects perform equally well on all three item forms.

◆ ◆ ◆

Type and Amount of Available Past Instances in Concept Learning.

DANIEL D. BLAINE, J. L. DUNHAM, THOMAS W. PYLE, The University of Texas at Austin

Several experiments have shown that making past instances available has a facilitative effect on performance. Although these results are generally interpreted as due to a reduction in memory load, there is still a question as to what it is about past instances that accounts for the superior performance under conditions of availability. A recent study has demonstrated that different types of availability differentially affect performance. In this study, it was hypothesized that available past instances facilitate performance to the extent that they provide information which is relevant to the solution of the concept learning task.

Fifty-nine high school students were given a memory for associations test. The distribution of scores was dichotomized at the median, in order randomly to assign *Ss* to conditions on the basis of their memory score. The concept learning task was a 6-dimension, 4-category problem with two binary dimensions relevant. The instances which *S* classified incorrectly were either made available in the correct category or the incorrect category which *S* chose with either 1, 2, or 4 instances allowed to accumulate in each category. In addition, a control group was run in which no instances were available.

The availability of past instances in the correct category facilitated the performance to a significantly greater extent than instances in the incorrect category. The major effect of amount of availability was found between zero instances available and one instance per correct category available with the facilitative effect leveling off as the amount of availability increased. Memory was found to interact with the type of availability but not the amount of availability. It was concluded that the facilitative effect of available instances occurs only insofar as the instances provide information regarding the correct categorization of instances.

9:00 to 9:55 A.M. SESSION 14.6 DIVISION C

PROGRAMED INSTRUCTION III

CHAIRMAN:

Harry Schumer, University of Massachusetts

PAPERS:

Individual Differences and Task Parameters in Learning by Computer-Assisted Instruction

HENRY T. LIPPERT, Florida State University

The SOCRATES computer-based teaching system was used to study the effects of four different teaching strategies on the learning, retention, and transfer of skills in pronunciation of Russian words. The four macro-teaching strategies used differed in the amount of prompting and cue-support offered at different thirds of practice. The results indicated that significant amounts of learning retention and transfer were obtained.

Students thought they would prefer to learn words six at a time before learning, but changed to a preference of twelve words at a time after learning. The student's confidence in his level of learning increased as learning took place. Confidence attained its highest levels when three words were taught at a time and when a teaching strategy was used in which cue-support was available, then withdrawn, and then made available again. Students with high aptitude were more confident about their performance, but did not obtain better scores on the performance tests than did low aptitude students. Most of the time taken to learn (latency) was during the micro-strategy (prompting) used first in all four macro-teaching strategies. Significant differences in performance were not obtained in reference to the four macro-strategies of teaching used. The inclusion of a few prompting S-R trials had a potent effect. Latency was not significantly related to aptitude. The latency data indicated that the actual time taken to learn the words was independent of how many words were presented at a time. Carroll's model (1963), based strongly on time factors, was not supported by the results obtained. For the task used, the student's confidence level was much more strongly related to aptitude and performance than was his response latency.

◆ ◆ ◆

A Computerized Determination of the Readability of Programmed Materials Using Complete Units

ROBERT A. SHAW and MILTON D. JACOBSON, University of Virginia

The purposes of this study were: (1) to automate a procedure which could be used to obtain total variable counts and complete word lists;

and (2) to produce a regression equation which could be used to determine the reading difficulty of the programmed materials.

The sample of materials consisted of three units, and the pupil sample consisted of 559 elementary school children in grades four, five, and six.

An error count was used as the measure of the criterion variable, and the variable was validated by correlating the counts on this variable with five measures of achievement and an IQ score. The correlation coefficients were all negative (as predicted) and significant at the 0.01 level.

The independent variables were determined by an analysis of the literature and the programmed science materials. Counts on these variables were obtained by using a computer. Seventeen independent variables were chosen to be used in the initial regression equation to predict the dependent variable. A regression routine for the computer was used to obtain the equations from the sampling procedures, and deletions were made to determine which independent variables contributed significantly to the final equation. Null hypotheses were stated and checked for each of the seventeen independent variables, and for the significance of difference between the multiple R's using different sampling procedures: (1) random sampling; (2) sequential sampling ; (3) cluster-sequential sampling; and (4) total sampling, on the established mathematical regression models.

Only three of these variables were found to contribute significantly to the regression equation for predicting the criterion variable: (1) the average number of mathematical and scientific words per sample; (2) the per cent of frames that were response frames per sample; and (3) the average number of review frames per sample, respectively. The final regression equation had a multiple R of 0.502.

◆ ◆ ◆

Effects on Achievement and Attitudes of Two Writing Styles Used With Programmed Materials

THOMAS S. NAGEL, Michigan State University

The objectives of this study were to demonstrate that a conversational and anecdotal style of writing when compared with a formal style of writing used in programming the "Elementary Principles of Celestial Navigation" would result in significantly greater achievement and better attitudes toward the program and toward programmed instruction in general.

Two forms of the linear program were developed, differing only in styles of writing. The programs were identical frame by frame for subject content, and both were 81 frames long. Eighty naval reservists in pay grades E3 and below were randomly assigned the programs (38 formal and 42 conversational), and completed a posttest and an attitude questionnaire.

Results of the posttest using a 2 × 2 analysis of variance where the two forms of programs were contrasted with prior experience with programmed instruction (or lack of it), showed that there was a significant treatment effect at the .01 level. A second 2 × 2 analysis using experience celestial navigation, instead of experience with programmed instruction, resulted in a significant treatment effect at the .001 level. It was concluded that the type of writing used had a significant effect on the performance of individuals who had had no prior experience with programmed instruction or with the subject matter being studied.

Results of the attitude questionnaire were inconclusive. A chi square analysis of the data showed that five of the 24 questions were significant, but no generalizations could be made.

The important result of the study was that for maximum achievement it appears to be advisable to use a conversational and anecdotal style of writing when it is expected that the users of the program may never have worked with programmed instruction before, or are ignorant of the subject matter contained in the program.

◆ ◆ ◆

An Evaluation of the Efficiency of the Instructional Program, Effective Listening

H. W. GUSTAFSON, Bell Telephone Laboratories, Inc., and HARRY A. SHOEMAKER, American Telephone & Telegraph Co., (Ernest Z. Rothkopf, Bell Telephone Laboratories)

Prospective purchasers of off-the-shelf instructional programs rarely possess the wherewithal to assess the absolute effectiveness of the programs before deciding whether to buy. In most instances, however, they have at their disposal a very simple means of investigating the question of program efficiency—i.e., the cost of a program in comparison to alternative ways of accomplishing the same job. It consists of opposing a given program to an alternate version of itself designed either to shorten learning time or to reduce manufacturing costs. It is easy to shorten the study time for almost all linear programs, e.g., merely by filling in the answer blanks beforehand. Other approaches to program abbreviation include (a) reorganizing the content into simple lists and tables; (b) summarizing the program in the form of a narrative, and, in the case of branching programs; (c) eliminating all diagnostic questions and remedial material.

Two studies of this sort were done to evaluate the efficiency of the two and one-half hour, tape-recorded program, *Effective Listening*. The first compared the program (N = 18) to roughly 15 minutes spent reading a summary of its content (N = 16). The second contrasted *Effective Listening* (N = 23) to both a no-training control treatment (N = 21) and a tape-recorded revision of the previous summary involving approximately 30 minutes of training (N = 21). The subjects were female employees of the Bell System. The pretest and posttest items consisted of tape-recorded statements which the subjects listened to and summarized in writing.

In the first study *Effective Listening* yielded significantly larger gains than the printed summary. In the second the program also proved better than the summary (P = .06), but only with respect to sanctioned test items furnished by the vendor. With respect to other items of similar nature constructed or selected by the investigators, the program was not superior to the summary nor even to the no-training control treatment. The latter finding would appear to cast doubt on the generality of the skills taught by the program.

9:00 to 9:55 A.M. **SESSION 14.8** **DIVISION C**

VERBAL BEHAVIOR V

CHAIRMAN:

Julius M. Sassenrath, University of California at Davis

PAPERS:

Patterns of Parent-Child and Teacher-Pupil Verbal Interaction: Implications for Training

ELIZABETH HUNTER, Hunter College, New York

This paper compares the findings of the growing number of studies of parent-child verbal interaction (i.e. Bernstein, Henry, Shipman, and Hess) with the findings of the equally increasingly available studies of teacher-pupil verbal interaction (i.e. Bellack, Flanders, Gallagher, and Aschner, Hughes) in terms of the effect of parent and teacher language upon children's cognitive abilities and school achievement. In addition, means for training teachers and parents so that they may increase their understanding of the importance of language, as well as their range of verbal behaviors, (Amidon and Hunter, Ginott) are explored.

A search of the literature was made; both the literature dealing with the effects upon children of varying kinds of parent and teacher language, and the literature describing means of training for changed verbal behavior. The author is actively engaged in training teachers and future teachers to widen their verbal repertoires through the use of various verbal interaction category systems, and through skill sessions which involve role playing, written cases, and tapes of classroom talk. The paper reports the results of this work.

Studies of both parent and teacher verbal interaction with children indicate that language styles influence children's cognitive styles and learning abilities. Training results in increased awareness of the importance of language and the achievement of wider verbal repertoires.

Teacher educators are beginning to recognize the importance of teacher talk, and an increasing number of teacher education programs include the study of the language of the classroom. Little, of anything, is done to work with parents, in any organized fashion, to increase their awareness of the important consequences of their talk.

◆ ◆ ◆

The Effect of Reading for a Creative Purpose on Student Attitudes Toward a Short Story

RICHARD SMITH, Madison Public Schools (Wayne Otto, University of Wisconsin)

This study was designed to investigate the effects of using specific kinds of writing tasks, assigned prereading, and completed postreading, on student attitudes toward a short story. The following hypothesis was investigated:

If capable high school seniors are oriented to the reading of a short story by a pre-assigned, creative writing task, their attitudes toward the story will be more positive than if they are oriented to it by a pre-assigned, noncreative writing task.

Clarification of the effects of a pretreatment training program which taught basic differences between creative and noncreative writing tasks and written products on the attitudes developed was an ancillary concern.

Twenty twelfth-grade college-preparatory English classes from two Madison, Wisconsin, senior high schools judged to be alike in any variables likely to influence the results of the study were randomly assigned to four different treatments. Those assigned to treatment one received the creative writing task and the training program; those assigned to treatment two received the noncreative writing task and the training program; those assigned to treatment three received the creative writing task, but no training; and those assigned to treatment four received the noncreative writing task, but no training.

The experiment took one fifty-minute class period and proceeded as follows:

1. Study the writing task
2. Read the story
3. Complete the writing task
4. Complete the attitudinal inventory

The attitudinal inventories were scored and analyzed by the Reciprocal Averages Program. This computer program indicated a reliability of .88 for that instrument.

The mean class scores were subjected to a two-way analysis of variance to determine any significant differences among the four cells. A significant difference at the .05 level was found in favor of the noncreative writing task variable. This was in the opposite direction to that expected. No significant difference was found between the trained and untrained groups.

◆ ◆ ◆

A Study of the Relationship Between Perceptual Malfunction and Reading Achievement

WALTER A. BUSBY and DONALD E. HURD, Bradley University

Research attention to reading behavior in recent years has encompassed a variety of perceptual abilities. However, influence of perceptual malfunction has been investigated mostly through singular channel assessment. Achievement in reading could be hampered by perceptual difficulties in the processing of sequential information in both the visual and auditory sense modalities, while no defect would exist singularly. The purpose of this study was, therefore, to determine the relationship between visual and auditory perceptual processing of sequential information to reading achievement with control of the intervening variables of intelligence, age, sex, and socio-economic status.

Subjects were selected from the 38 public elementary schools in Peoria, Illinois. In all, 120 students were chosen at random from grades two, four, and six. Measures of the following variables were collected on each student in the sample:

1. Singular visual-auditory perception;
2. Visual auditory shifting perception;
3. Intelligence (Wechsler Intelligence Scale for Children);
4. Socio-economic status (Socio-Economic Scale);
5. Reading achievement (California Achievement Test—Reading).

The perceptual apparatus used in this experimentation has been described earlier by Sutton, et al, and differs only in equipment design. The stimuli consisted of a red light, green light, and low and high tones of 400 cps and 1200 cps. Presentation was automatic with random interval timing of 1.5, 2.0, and 3.0 seconds. Mean reaction times were then computed for (a) stimuli preceded by a stimuli in the same channel and (b) stimuli preceded by stimuli in a different channel.

The present study rejected the hypothesis that auditory and visual perceptual processing of sequential information is related to reaching achievement. The conclusion appears to contradict previous research findings. However, these earlier studies were primarily concerned with differences between high and low reading ability groups. Their findings suggest that low and high reading ability groups differ significantly in perceptual ability. Hence, while perceptual malfunction may be a factor in low reading ability samples, the present study sugggests that it is not a factor in a representative sample.

◆ ◆ ◆

Effect of Time of Transition from Manuscript to Cursive Writing upon Subsequent Performance in Handwriting, Spelling, and Reading

WAYNE OTTO and G. LAWRENCE RARICK, University of Wisconsin

There is support for the initial introduction of manuscript style handwriting, but there is neither consensus nor definitive research regarding the most advantageous time to make the transition from manuscript to cursive writing. The general purpose of this study was to determine whether there is an optimum time for the transition. Specifically, the study was designed to determine whether the handwriting, reading, and spelling performance of fourth and sixth grade children is related to the time at which the transition from manuscript to cursive handwriting was made.

The State of Wisconsin was surveyed to identify school systems that employ the four most common transition times: Fall, Second Grade; Spring, Second Grade; Fall, Third Grade; Spring, Third Grade. From the systems identified, three were chosen at random to represent each transition time. In turn, five boys and five girls were selected at random from grade four and from grade six of each system to participate in the study. Both fourth and sixth graders were included in order to get evidence relevant to the notion that perhaps relationships would change with the passage of time and interjected experience. Thus, there were 240 subjects, chosen according to sex, grade level, and transition time.

The reading and spelling subtests of the California achievement battery were administered to the subjects and handwriting samples were obtained. Raw scores from the California subtests, letters produced in one minute of writing and legibility rating scores, based on the ratings of three trained judges using two different scales, were examined by analyses of variance and appropriate post hoc tests.

There was no support for the notion that time of transition may affect subsequent reading performance, and only tangential evidence of a relationship between transition time and subsequent spelling performance. Time of transition apparently did affect subsequent handwriting performance: late transition was associated with rapid writing and early transition with legible writing.

9:00 to 9:55 A.M. **SESSION 14.9** **DIVISION D**

RESEARCH ON TESTING III

CHAIRMAN:
Kenneth F. McLaughlin, U.S. Office of Education

PAPERS:

Computer-Based Conversational Score Reporting

CARL E. HELM, City University of New York (Albert J. Harris)

The *effective* reporting of test results to admissions officers, guidance counselors, and the individuals who take educational tests requires the communication of the statistical, psychological, and educational contexts which allow the user to infer relevant meanings. Educational testing services prepare a wide range of interpretive materials for most testing programs to aid users in interpreting scores. Since these materials typically cover a range of scores and content areas, a relatively sophisticated analysis by a user is required in order to arrive at those interpretations that are relevant to the particular score profile under consideration.

A computer-based system has been designed to write letters interpreting score profiles using existing interpretive materials as a basis for the content of the letters.

There are three major facets to the problem of designing such a system: 1) There must be a simple but flexible set of procedures for specifying the verbal interpretive materials appropriate for specific score profiles; 2) an efficient computer program must be prepared for the actual processing of materials; 3) a complex distribution system must be perfected for the dissemination of information to individuals. In the present research we have addressed ourselves to the first two facets of the problem. On the third facet, problems of privacy and accuracy of a different order of magnitude than those presently encountered must be solved.

◆ ◆ ◆

An Investigation of the Validity of the Order of Items on the Peabody Picture Vocabulary Test

JOSEPH S. RENZULLI and DIETER H. PAULUS, Bureau of Educational Research University of Connecticut

The Peabody Picture Vocabulary Test (PPVT) is a widely used instrument designed to measure the intelligence of English speaking residents of the United States. A subject's score on this test depends upon how many items are correctly answered by the subject, however, testing is stopped if the subject misses six out of eight consecutive items. Therefore, a necessary condition for the validity of this instrument is that the items be arranged in ascending difficulty so that if testing is stopped, the subject would in fact, not answer any more items correctly if the test were to be continued.

It was the purpose of this investigation to empirically evaluate the validity of the ordering of the items on the published version of the Peabody Picture Vocabulary Test. In order to do so each form of the PPVT was administered to 450 subjects. Item difficulty indices were calculated for each item on the test and items were re-ordered on the basis of these indices. A rank-order correlation coefficient was calculated between this new ordering and the published ordering of items. This step of the research indicated that the ordering of the items could be substantially improved. In order to do so, 225 answer sheets were randomly selected from the original sample and again item difficulty indices were calculated. Items were then re-ordered on the basis of these indices and the new ordering was validated on the remaining 225 tests. Again a rank-order correlation coefficient was calculated to evaluate this new, cross-validated ordering of items on the PPVT.

The original rank-order correlation was compared to the one obtained in the cross validation sample and it was found that the validity of the item order on the PPVT was improved. Statistical evidence reporting the size and the significance of the increase in validity is reported.

◆ ◆ ◆

Effects of Promised Reward and Penalty on Performance of a Multiple-Choice Vocabulary Test

ROSS E. TRAUB, R. K. HAMBLETON and B. SINGH, The Ontario Institute for Studies in Education

In multiple-choice tests students may be told to guess when the answers to questions are not known, or they may be told *not* to guess, or they may be told nothing. When the instructions say not to guess, motivation for obeying is usually provided by additional instructions threatening application of a penalty for wrong answers. There is another way to induce the examinee to obey instructions not to guess—a way that has been ignored in previously reported research: It is to promise reward for omitted items. The purpose of the present investigation was to compare the effects on performance of a multiple-choice vocabulary test of promising reward and threatening penalty in instructions not to guess. Base-lines for comparison purposes were provided by a) instructions to guess whenever in doubt, and b) no instructions at all about guessing.

The study was implemented by randomly dividing 667 ninth-grade students into four groups. Each group read a different set of guessing instructions and then worked, under power conditions, two (essentially) parallel forms of a 90-item vocabulary test. The test was chosen to be difficult for the students in order to provide ample opportunity for guessing.

Instructions *not* to guess under promise of reward resulted in dramatically (and significantly) fewer wrong answers and more omitted questions than the other conditions. These same instructions were associated with a mean number of correct answers that was not significantly different from the mean associated with instructions *not* to guess under threat of penalty. Both of these means were significantly lower than those observed for the groups receiving the control instructions. Finally, the parallel-forms estimates of reliability were significantly higher for two groups who themselves did not differ: the one instructed *not* to guess under promise of reward, the other instructed to guess whenever in doubt. No systematic differences attributable to guessing instructions could be observed in the Kuder-Richardson estimates of reliability.

◆ ◆ ◆

Information Removed from Multiple-choice Item Responses by Black-out of Selected Grammatical Categories

WENDELL W. WEAVER, The University of Georgia, and A. C. BICKLEY, Campbell College

This study involves blacking out certain grammatical classes, e.g., nouns, and studying the effect of the black-out on the number of correct responses to test items concerning that passage.

There are seven treatments: nouns blacked-out; main verbs blacked-out; adjectives blacked-out; nouns, main verbs, and adjectives blacked-out; all words blacked-out except nouns, main verbs, and adjectives; all words blacked-out; and no words blacked-out.

Reading paragraphs were selected at random from reading tests listed in Buros' Sixth Mental Measurements Yearbook. Selection was limited to expository paragraphs. All items connected with a particular paragraph were kept intact and in the same order as in the original test.

Subjects were given the Vocabulary subtest of the WAIS and ranked by total score. The entire group of vocabulary scores were divided by thirds into a high, middle, and low group. Seven subjects from each third of the vocabulary ranking were randomly assigned to each of the seven treatments. There were 147 subjects in all, 21 Ss per treatment, 7 high, 7 middle, 7 low, on vocabulary scores.

A treatments X blocks ANOVA was computed. There was a statistically significant treatment difference at the chosen (.05) level of confidence.

It was found that most information was lost, in the case of single grammatical categories, when nouns were deleted. Deleting main verbs resulted in about the same loss of information as when adjectives were deleted. The test items presented alone supplied a large amount of information for their completion. As noted in other studies, the most deleterious condition was the blacking-out of nouns, verbs, and adjectives, thus leaving mostly function words. Under this condition significantly more information was lost in responding to test items than under the "no paragraph" treatment.

9:00 to 9:55 A.M. **SESSION 14.10** **DIVISION D**

CONTRIBUTED PAPERS II

CHAIRMAN:

Robert G. Stakenas, Florida State University

PAPERS:

Teacher Tasks and Participant Observation: Strategies for the Measurement of Teaching Utilizing Tasks, Films and Observers

FRANK GARFUNKEL, Boston University

Observational records used in comparisons of teaching are subjected to methodological control and conditions of reversibility by filming predetermined tasks that are carried out in classrooms by teachers. Films are the data that are abstracted from teaching situations. Scales have been developed to further abstract dimensions of differences between teachers that are irreducibly presented by films and by participant observer reports.

Objectivity is a quality that must be applied to not only scoring procedures, but to substance and reactive effects. The use of "bit" ratings or "bit" behavioral recordings can be reliable while, at the same time, of questionable validity. It is not at all clear what the optimal reduction strategy is, but only that objectivity will be suspect if measurement of teaching is not handled in diverse ways. Prediction can aim at specific bits of behavior over an extended period of time, traits that characterize a teacher in various situations or, as in the case of the reported study, by audio and visual recordings of teaching behaviors in a limited number of situations, but with relatively complete detail. Comparative judgments can best be made by skilled (as observers) and experienced (as teachers) observers who have adequate opportunities to observe classes, interview teachers and be fitted into a design which permits replication and multiple observation. Films can be used to anchor judgments and to insure that the sample of classes for any one observer is sufficiently variable and representative to justify comparable scale distinctions to any other observers.

Procedures and data from the three observational studies will be reported, two of which include films in the presentation of data. The first study is a comparison of inner city and suburban classrooms and is a part of a larger study of the effects of bussing inner city children to suburban schools. Eight classrooms (kindergarten through five) have been filmed and 40 observed and judged by participant observers, on a series of specifically designed scales. The second study involved the filming of six Head Start teachers, each carrying out three tasks. The third study involved a participant observational comparison of twenty Head Start classes which were also rated by other procedures.

The paper consists of a rationale, elaboration of the model, alternative procedures, and results in the form of quantitative comparisons, correlations between concurrent measures and a selection of filmed comparisons.

Word Frequency of Association Chains in Children as a Function of Word Frequency of the Stimulus and the Chain Length

PAUL OWOC, ROGER BISHOP, and THOMAS J. JOHNSON,
Washington University

Verbal mediation is usually conceived of as a chain of covert associative responses which function between an overt initiating stimulus and an overt terminating response. This study began as an attempt to develop a methodology for generating chains of associations in order to investigate the influence of certain constraints on their characteristics. The methodology of investigation was an integrative association task in which subjects were given a beginning and an ending stimulus word and were instructed to provide other words which would serve to connect the two stimulus words in a meaningful logical manner. These interpolated words are referred to as integrative association chains. The purpose of this study was to investigate the influence of two variables, chain length, and word frequency of the stimulus words, on the word frequency of the chain of associations.

a) There was an overall tendency to form associations which consisted of words which were of relatively high frequency of occurrence in the Adult Vocabulary; b) The mean word frequency of the associations tended to be higher when the frequency of the first stimulus word was of low frequency ($p < .01$; Table 1); c) More chains were generated when the frequency of the first stimulus word was high ($p < .05$); d) Seventh graders tended to rate the task as being more difficult ($p < .05$); and e) Chains of moderate length (4 or 5 links) are rated as more difficult to construct than are long or short chains ($p < .05$).

The discussion will focus on the nature of the associative process in children.

◆ ◆ ◆

The Relation of Dogmatism to the Complexity of Individual Semantic Meaning Spaces

KEN SIROTNIK, University of California, Los Angeles
(T. R. Husek)

The purpose of this study was to investigate the nature of the dogmatic personality via the analysis of semantic differential responses. The highly dogmatic individual was hypothesized to be one who is rigid, intolerant of ambiguity, and unable to make judgments in gradations; the person low in dogmatism was conceptualized as demonstrating a polar opposite behavioral pattern. It was predicted that given positive and negative authority figures as concepts on a semantic differental, the individual low in dogmatism, by taking concept-scale interactions into account, would display a more complex semantic meaning space than the individual high in dogmatism. Furthermore, the high dogmatic would tend to choose scale point extremes more often than the low dogmatic.

The data used in this research were obtained from another study. Using Rokeach's Dogmatism Scale to define extreme levels of dogmatism, 31 high and 27 low dogmatics were identified from a sample of 759 elementary school teachers and principals. Five measures of complexity were defined based on a principle components analysis of the intercorrelations of scales over concepts for the individual semantic differential. An additional measure was used to examine the hypothesis of scale point extremity. Specific predictions relating each of these measures to extreme levels of dogmatism were tested. None of these predictions was borne out.

The negative conclusion of no relationship between the major variables was examined in the light of the two instruments used in the study. The implications of the factorial complexity of the D-scale as perhaps obscuring the results are discussed as well as the possibility that the specific semantic differential instrument was inappropriate to examine the hypotheses under test. Implications of these findings for the future testing of the hypotheses are discussed.

◆ ◆ ◆

Word Associations of Blind and Sighted Children to Selected Form Classes

M. H. TILLMAN and CHARLOTTE WILLIAMS, University of Georgia

More than thirty years ago, Payne (1933) studied free associations of blind children to a modified version of the Kent-Rosanoff word list. She reported that blind children gave more idiosyncratic responses, more perseverative responses, and fewer common responses; consequently, Payne suggested that blind children had a restricted field of association. The present investigation, also using word association techniques, was designed to examine levels of syntactic development in blind and sighted children. Since association on the basis of syntactic function, that is, homogeneous or paradigmatic responding, is highly related to the development of syntax in children, word association responses in the present study were scored correct when the response member was the same part of speech as the stimulus word.

Thirty-five blind children, seven at each age level from seven through eleven years, were obtained from three Southeastern residential schools. A sighted group comparable to the blind in age, intelligence, number, and socio-economic background was obtained from three communities in Georgia. Two tasks were involved: a word association (WA) task (Brown and Berko, 1960) consisting of 36 high-frequency stimulus words from six parts-of-speech categories, count nouns, mass nouns, transitive verbs, in-transitive verbs, adjectives, adverbs; a word usage (WU) task consisting of nonsense syllables introduced within a syntactical context such that the six form classes of interest were represented. In the latter task, the child was asked to supply a meaning for the nonsense syllable.

Results (Groups × Levels Anova) indicated that on WA tasks both groups scored about the same level, that rate of development was about the same, and that the sequence of development was highly correlated. On WU, however, a ceiling effect was obtained for the blind group but not the sighted group. Several possible explanations were considered: biased low estimate of the intelligence of the blind group; effects of heterogeneity of variance; a possible confounding of subject and task variables; and, a face-value hypothesis. It was suggested that the use of younger children would clarify the nature of differences obtained on the WU task. In short, while it is possible to infer specific language deficiencies based on research literature, the results provided no support for a developmental lag in the blind group's acquisition of syntax.

10:00 to 10:55 A.M. SESSION 15.1 DIVISION C

PROBLEM SOLVING AND THINKING

CHAIRMAN:

Douglas D. Sjogren, Colorado State University

PAPERS:

The Relationship Between Parental Behavior and the Development of Flexible Thinking in Boys

THOMAS V. BUSSE, University of Chicago and The Institute for Juvenile Research

This study explored the relationship between parental behavior and the development of boys' flexible thinking, defined as the ability to consider alternative means to a given end. The subjects were 48 Negro fifth grade boys and their natural mothers and fathers from a mostly lower-class, semi-rural, Negro community.

Twelve variables tapping both verbal and nonverbal components of parental behavior were scored from a session in which each parent separately taught the child four tasks. The boys' flexible thinking scores were obtained from a factor analysis of six flexible thinking tests (e.g., Match Problems, Embedded Figures Test, Water Jug Test) and eight reference tests.

Linear and quadratic trend analyses were computed using the various parent behaviors as independent variables, and the boys' flexible thinking scores as the dependent variable.

The results showed that both mothers and fathers who manipulated the task materials a moderate amount, had sons who were higher in flexible thinking than did mothers and fathers who manipulated the materials either very little or very much. For both mothers and fathers, high manipulation scores appeared to hinder the development of flexible thinking even more than low manipulation scores. In addition, linear relationships were found linking mother commanding behavior, father total words, and father love with the boys' flexible thinking. Mother pointing also showed a quadratic relationship to flexible thinking.

◆ ◆ ◆

An Examination of Selected Critical Thinking Behaviors of Children in Grades Five Through Nine

HENRY B. SLOTNICK, University of Illinois

This study examined the abilities of boys and girls at the indicated grade levels to (1) decide which of several conclusions followed most reasonably from one or more major and minor premises (conclusions behavior); and (2) which of several assumptions could most reasonably have been made for one or more assumptions (assumptions behavior). Each behavior was measured at three levels of sophistication:

1. one major premise (conclusions behavior) or one statement (assumptions behavior);

2. a multiplicity of related major premises or statements. Levels 1 and 2 utilized content with which (it was assumed) the child was familiar;

3. The content the child was asked to handle was abstract; nonsense words were used.

A pencil and paper, machine scored instrument was administered to 392 students. The effects of grade and sex were evaluated using a fixed effects multivariate ANOVA where $X_{1,8}$ was the i^{th} individual's scores on the subscales of the test.

The ability of children of any given sex and grade to perform either of the described critical thinking behaviors at a given level of sophistication appears to be related to both the behavior in question and the level of sophistication.

Critical thinking ability appears to vary with grade. The discriminate centroids indicate two clusters: the first is composed of eighth and ninth graders; the second contains fifth, sixth, and seventh graders. The scaled vector associated with this discriminate function weights the abstract subscales heavily relative to the familiar subscales.

Critical thinking ability appears to be related to sex. However, no discriminate function was found that distinctly separated the centroids for boys' and girls' groups.

Considering the critical thinking behaviors in terms of independent and dependent variables, the difficulty of the critical thinking situation increases more rapidly with the number of independent variables than dependent variables.

◆ ◆ ◆

Logical Thinking in Children

BERNARD J. SHAPIRO, Boston University and THOMAS C. O'BRIEN Educational Research Council of America

Research done by Hill, widely reported by Suppes and replicated by Osborne, has provided evidence, contrary to Piaget's theories, that children at age six, seven, and eight are capable of hypothetico-deductive thinking and are by no means limited to "concrete" thought operations. In both of these studies logical thinking was measured by a 100-item test consisting of three major categories in elementary logic—sentential logic, hypothetical syllogism, and logic of quantification. Each item consisted of two premises and a question that required the student to discriminate between a necessary conclusion and the negation of that necessary conclusion. In both studies children scored higher than would be predicted on the basis of Piaget's theories.

The present investigation was undertaken to measure children's ability to test the logical necessity of a conclusion and to determine whether differences existed between this ability and the originally tested ability. Therefore, the present study was undertaken in two parts: the first was a direct replication of the original experiment; while the second used virtually the same instrument with some items modified so that no necessary conclusion followed from the premises.

Analyses of covariance, controlling for IQ, supported the original research in that for the original test similarly high scores were obtained, no differences were detected between sexes, and statistically significant differences occurred between age groups.

On the other hand, the analyses also indicated that for all age groups the modified instrument was much more difficult than the original. Statistically significant differences of substantial magnitude were detected between children's performances on the two tests at each of the age levels. Interestingly, not only were there no differences between sexes on the modified instrument, but there were also no statistically significant differences between the six, seven, and eight-year-old groups on the modified test.

On the basis of this investigation, it is questioned whether available evidence contradicts Piaget's theory with respect to hypothetico-deductive thought in children at age six, seven, and eight.

◆ ◆ ◆

Observation vs. Participation in Learning to Discover

ROBERT C. CRAIG, Michigan State University

Gagne and Brown (*J. Exp. Psychol.*, 1961) showed that a small-step program may be used to teach the discovery of rules (formulae) for the sums of number series. The present experiment used that same program in a search for evidence that students can learn discovery strategies by watching others, as well as directly.

Subjects of two samples, high school boys and college students, were randomly divided to form a control and three experimental groups. Each subject of a *Direct Discovery Group* worked through a programmed booklet in the conventional manner. The *Student Observation Group* watched while fellow students worked through the program aloud with the teacher.

The *Teacher Observation Group* watched and listened while the teacher asked and answered each question of the program himself. The *Control Group* studied a programmed introduction to programmed instruction.

All treatments were repeated on two different days. Then, a test on the discovery of rules for new number series was given. The treatment groups were compared on time to discover the new rules, number of hints needed, and a weighted time score combining time and number of hints. The differences among the means of the three experimental groups on these criteria were not significant for either the high school or the college sample, but all experimental groups required less time and fewer hints than control groups.

The evidence of this study is consistent with the suggestion that students can learn discovery vicariously in groups, as well as directly and individually. In programs to teach discovery, as in others, an overt response is not always essential. The methods of observation used here are somewhat unnatural and time consuming; but they have the merit of being describable and repeatable, and could provide a point of reference for the systematic evaluation of other proposals for group instruction.

10:00 to 10:55 A.M. **SESSION 15.2** **DIVISION C**

REINFORCEMENT II

CHAIRMAN:

Douglas Porter, Harvard University

PAPERS:

The Effects of Teachers' Reinforcement Style Upon Imitative Behavior of Children

STEPHEN H. PORTUGES and NORMA D. FESHBACH, University of California, Los Angeles

This study is designed to provide information concerned with the indirect multiple effects of teachers upon social behaviors of particular kinds of children. The specific intent of this investigation is to evaluate the effects of variations in teachers' reinforcement style on imitative behavior in children varying in personality characteristics and social background.

The following hypotheses were made:

1. advantaged children will exhibit more imitative behavior than disadvantaged children;
2. within each social class girls will manifest significantly more imitation;
3. there will be more imitation of the positive teacher;
4. a positive correlation between dependency and imitation will be obtained in the advantaged group.

48 disadvantaged and 48 advantaged boys and girls aged eight to ten years saw two films, each of which depicted a different teacher presenting a geography lesson on Africa. Teachers displayed distinct gestures incidental to the lesson, the imitation of which constituted the main dependent variable. The teachers also differentially reinforced answers from the children in the films; the positive teacher responding with approval to correct answers, the negative teacher emitting negative remarks to incorrect replies. After viewing both films, the subjects were given the opportunity to "teach" the geography lesson to two life-sized dolls. The identical procedure for the imitation sequence in which the children were instructed to teach the lesson was carried out with the 24 Control subjects who did not view the films. Measures of the children's dependency, aggression, and cognitive style were also obtained.

The four experimental hypotheses were statistically confirmed and interpreted in terms of the perceived similarity between teacher and child, and secondary reinforcing properties of the teacher. The implications of the absence of imitative behavior in the disadvantaged boys are discussed in relation to the teaching of the disadvantaged child.

Behavior Modification and Transfer in an Out-Of-Control Classroom

MICHAEL E. J. ORME and RICHARD F. PURNELL, Harvard University

Inner city pupils in an out-of-control classroom were exposed to reinforcement and modeling procedures designed systematically to regain control, establish learning-oriented habits, and maximize teacher effectiveness. The out-of-control classroom was one in which markedly destructive, aggressive, and impulsive behaviors predominated (e.g., fighting, screaming, running about the school, destroying materials and desks). Only stringent disciplinary reactions by school personnel were capable of restoring even short-term order. Achievement of educational goals was hopelessly minimal.

A class of 18 third and fourth grade children was divided into two sections (rooms). Following baseline observation, Section (B) pupils were continuously positively reinforced with consumable and verbal rewards for displaying stipulated educationally relevant behaviors. Desirable behaviors were reinforced by tokens (points on the blackboard) which could only be earned in Room B, and which could be exchanged for items such as candy, comics, field trips, preferred school project, or individual instruction in an area of interest. Students in section B were drawn from the previous day's section A, selection being based on points accrued for manifesting desirable behavior in that room. This pupil had to "earn" a place in Room B where they could then work for tokens. Simultaneous with establishing control, teaching techniques calculated to elicit and immediately reinforce active student participation in lessons were implemented in Room B by a teacher trained in the use of these techniques. Room A had a regular teacher. Six randomly scheduled five-minute intervals of total classroom behavior were video-taped from each section's daily three-hour periods for five weeks. Selected teacher and pupil behaviors were reliably rated and tallied from these tapes.

Dramatically significant reductions in disorderly classroom behaviors and increases in learning-participation behaviors were realized. The continuous turnover of pupils between Rooms A and B successfully effected a transfer of control to Room A.

◆ ◆ ◆

Verbal Satiation on a Positive and a Negative Reinforcer in Children

HAROLD COOK, Syracuse University

Satiation implies change in the meaning or effectiveness of a stimulus following its continued repetition. It was hypothesized: (1) repetition of a positive or negative verbal reinforcer would decrease its reinforcer effectiveness; (2) effectiveness of a positive or negative reinforcer would be altered when the repeated word was either positive or negative; (3) effectiveness would be enhanced following 10 repetitions, and diminished following 20, 30, 40, and 80 repetitions.

240 third grade children were randomly assigned in a 2× 2 × 2 × 6 fixed effects, factorial design. Factors were: (1) satiator, "good" or "bad"; (2) reinforcer, "good" or "bad"; (3) order of a semantic differential rating, either "before" or "after" conditioning; (4) amount of repetition, zero, 10, 20, 30, 40, or 80 repetitions.

Side preference was determined for each S in a free operant task involving dropping marbles into either of 2 holes for 50 trials. Satiation followed. Each S repeated aloud: "good" or "bad" for the designated number of repetitions. Half the S's then rated the word they repeated and the word "candy" on 3 semantic differential scales. Ss returned to the marble game for 50 conditioning trials and were verbally reinforced by E. "Good" was used by E for Ss in appropriate groups for marble drops Ss made to their low preference side. "Bad" was used by E for appropriate groups for marble drops Ss made to their high preference side. After conditioning, Ss rated the repeated word and "candy." Ss returned to the marble game for 50 extinction trials, but were not reinforced. Subsequent to extinction, Ss rated the repeated word and "candy."

There were two sets of dependent measures:

(1) task performance measures: (a) number of responses made to the low preference side during baseline, conditioning, and extinction; (b) response shift, from baseline to conditioning, and from baseline to extinction; (c) playing time, time from trial 1 to 50 during baseline, conditioning, and extinction; (d) playing time shift, from baseline to conditioning and from baseline to extinction;

(2) semantic differential rating measures: (a) polarity scores of the repeated word for Ss who made 2 ratings; (b) for Ss who made 3 ratings; (c) polarity scores of "candy" for Ss who made 2 ratings; (d) for Ss who made 3 ratings.

Results indicate that continual repetition of a verbal reinforcer influences the reinforcer effectiveness in a curvilinear manner ($p < .01$). Ss who repeated the reinforcer for 20, 30, or 40 repetitions evidenced a decrement in renforcer effectiveness below that of the zero or 10 repetition groups ($p < .01$). There were no statistical differences between zero and 10 repetitions. 80 repetitions Ss demonstrated an increase in reinforcer effectiveness above that obtained for 20, 30, or 40 repetitions. No statistically significant differences occurred across amounts of repetitions on the semantic differential data.

The concept of the orienting response and its habituation was evoked as a suggestive explanation of the findings.

Token Reinforcement in Regular Public School Classrooms

R. J. KARRAKER, University of Missouri at Kansas City

Token reinforcement has been employed to initiate and maintain academic behaviors with retarded children (Birnbeauer, *et. al.*, 1965; Lent, 1966; Staats and Butterfield, W. H., 1965); dropouts (Clark, *et. al.*, 1966) low achievers (Wolf, *et. al.*, 1966); and juvenile delinquents (Cohen, *et. al.*, 1966). However, there is no published evidence of token reinforcement systems implemented in normal public school classrooms. This paper summarizes the results of four attempts of applying token reinforcement systems in normal classrooms, individually none of which is publishable, but which collectively provide evidence that such systems can be useful.

The four teachers who implemented the system were regularly employed classroom teachers who taught from 24 to 35 students from grades K through 10. The teachers used tokens to reinforce successfully certain individual behaviors including turning in work on time, improving grades, fighting on playground, "talking out" in class, and studying quietly during study periods. Class tokens were also employed to modify such class behaviors as refraining from chewing gum, turning in work on time, returning report cards on time, and increasing achievement in spelling, arithmetic, and social studies.

The reinforcers backing up the tokens included trips into the community, opportunity to select activity, amount of time spent in class, parties on Fridays of each week, selection of traffic patrols, and opportunities to tutor younger children.

The individual behavior modifications were analyzed by an adaptation of the Fisher Exact Probability test; "t" tests were used for the class behaviors. Some of the modifications were run in a few weeks and others took several months. Problems with the system discussed were class size, administrative interference, difficulty in system maintenance in terms of recording and consistency, and inadequate teacher preparation prior to initiation of the system.

10:00 to 10:55 A.M. **SESSION 15.3** **DIVISION D**

RESEARCH ON TEACHING

CHAIRMAN:

P. Kenneth Morse, Medical College of Georgia

PAPERS:

Discriminate Analysis of Predictors for Career Commitment to Education

KATHRYN W. LINDEN and RONALD L. HOUSER, Purdue University

Each year an increasing number of teacher education graduates either fail to enter the profession or leave after a short involvement. This study extended a longitudinal investigation relating teacher attitude and personality characteristics to commitment to education, in this instance, six years in professional education. To date little research relevant to commitment to professional education has been reported. Both the criterion and the method of analysis employed in this study were thought to hold greater promise for research success than those employed heretofore.

During the spring of 1967, the 1961 teacher education graduates of Purdue University ($N = 322$) were asked, for the fourth time in nearly ten years, to report relevant life history status data and to complete two self-report inventories. Approximately 80% returned the life history data used to establish the criterion groups for this study. Five criterion groups were determined, ranging from six consecutive years in professional education to no entry into the profession. *Guilford-Zimmerman Temperament Survey and Minnesota Teacher Attitude Inventory* data obtained from these subjects upon entry into undergraduate teacher education were analyzed by multiple discriminant function procedures in order to determine which variates differentiated among the criterion groups.

Two of the four roots extracted exhibited discriminate efficiency. The variates in each root that significantly contributed to the group discrimination obtained suggested that humanistic, nurturant and status-seeking attitudes characterized those graduates who had persisted in educational careers. Both graduates who failed to persist and those currently teaching with three or less years of experience were characterized by materialistic and non-competitive attitudes.

Certain personality characteristics reported by beginning teacher education students differentiated career status six years after graduation. Measures of these characteristics should aid in the selection of candidates for teacher education who may be expected to fulfill their commitment to professional education.

Validation Results: A Performance Test of Teaching Proficiency

W. JAMES POPHAM, University of California, Los Angeles and EVA L. BAKER, Southwest Regional Laboratory for Educational Research and Development

The paper will describe results of a project undertaken to develop and, hopefully, validate a heretofore untried method of assessing teacher competence, namely, through the use of a performance test. Papers describing the rationale and interim results from the investigation have been presented at the 1966 and 1967 AERA annual meetings. This paper will describe the conclusion of the project, a final report of which is now available (USOE Project No. 5-0566-2-12-1, Final Report, August, 1967).

The performance test was designed to function in the following way. Teachers were presented with a list of specific, operationally defined objectives for a particular topic and directed to teach the objectives. Following the instructional period, students were tested on the behaviors stated in the objectives. Teacher competence was judged on relationship to the way their students performed on the criterion test. An attempt to validate this method of measuring teacher effectiveness involved contrasting the results produced by experienced teachers and nonteachers on a performance test dealing with research methods in the social sciences.

Two separate contrasts were conducted, the first involving six professionally trained, experienced student teachers versus six housewives for a six hour teaching period. The second involved 13 regularly credentialed teachers and 13 college students for a four hour teaching period. In neither contrast did the teachers perform significantly better than the nonteachers.

The results will be interpreted as indicating that the experienced teachers were not more experienced than the nonteachers in promoting leraner achievement of previously established instructional objectives. An alternative approach to validating the performance test strategy is discussed along with possible procedural modifications in the approach.

◆ ◆ ◆

The Effects of Training Observers of Classroom Behavior

JEANINNE N. WEBB and BOB BURTON BROWN, University of Florida

This paper reports the data obtained in a study of the effects of training in the use of an observational instrument for systematically describing the classroom behavior of teachers.

Thirty-six experienced teachers were trained in the use of the *Teacher Practices Observation Record* (TPOR) to achieve maximum observer reliability. The subjects of the study included 18 white and 18 Negro elementary teachers selected from a rural Florida county.

A brief description is given of the training procedures used; through the use of films trainees were provided with immediate feedback in order to increase between-observer agreement. However, the paper focuses primarily on the analysis of the data obtained in each of the five two-hour training sessions, including: (1) means and standard deviations of TPOR scores given by observers, (2) between-observer reliability coefficients, (3) within-observer reliability coefficients, and (4) frequencies of observer responses as a function of training.

The importance of this study lies in session-by-session report of what happens to reliability and validity of observations as the result of training. To our knowledge, this sort of analysis of observer training has never been reported before. Throughout the literature on observational systems observer training is reported as an accomplished fact without providing any data regarding the effects of training on observer responses. Likewise, such training is always done for the announced purpose of increasing reliability of the observations, and generally in terms of agreement between observers. These findings indicate that training has relatively little effect on observer-agreement *per se,* but has a startling effect on the validity of the observations with respect to the theoretical dimensions of the instrument and its sponsors.

◆ ◆ ◆

Biographical Characteristics of Effective and Ineffective Teachers

JOHN WILLIAMSON, Colorado State College

The purpose of this study was to determine whether biographical characteristics might be identified which would distinguish between groups of teachers receiving high and low student assessments. A second purpose was to analyze the students' reasons for assessing teachers high or low.

More than 2000 high school seniors were asked to identify the most effective and the least effective teacher they had experienced and to indicate why the teacher was effective or ineffective. This investigator personally administered a student questionnaire to the seniors in their schools. The student questionnaire utilized three major patterns of teacher classroom behavior identified by David Ryans. The three dimensions were:

1. friendly understanding behavior	*versus*	aloof, egocentric, restricted behavior;
2. responsible, businesslike, systematic behavior	*versus*	unplanned, slipshod behavior;
3. stimulating, imaginative, enthusiastic behavior	*versus*	dull, routine behavior.

Those teachers named by three or more seniors as either most effective or least effective were mailed a biographical inventory. Of the 257 teachers in both groups, 178 returned the completed inventory.

Analysis of data collected on 178 teachers and 2,200 high school seniors revealed the following results:

1. The mothers of ineffective teachers had significantly less formal education than mothers of effective teachers.

2. Effective teachers assumed most of the blame when a student was sent to the principal's office for discipline, whereas the ineffective teacher placed most of the blame on the student. This difference was statistically significant.

3. Ineffective teachers were more mobile than effective teachers.

4. In selections of effective teachers, seniors chose young (25-39) males and older (over 55) females.

5. Seventy-five per cent of the teachers named by the seniors were high school teachers.

6. The students' ratings were not related to the grade received from the teacher they rated.

Analysis of the seniors' comments regarding their teachers revealed the following information:

1. The effective teacher was genuinely concerned and interested in the students as individuals, required students to work, was impartial in dealing with students, and was obviously enthusiastic about teaching.

2. The ineffective teacher showed favoritism, was uninterested in students, was a poor disciplinarian, and obviously did not enjoy teaching.

10:00 to 10:55 A.M. SESSION 15.4 DIVISION D

EXPERIMENTAL DESIGN

CHAIRMAN:

Clinton I. Chase, Indiana University

PAPERS:

Error of Measurement and the Power of a Statistical Test

T. ANNE CLEARY, University of Wisconsin, and ROBERT L. LINN, Educational Testing Service

Educational researchers interested in comparing test results for groups of students are frequently faced with the problem of determining the most efficient research design within the constraints of available resources. A consideration of the relative influences of test length and sample size can be of considerable relevance for decisions about efficient design.

The purpose of this research was to study the effect of error of measurement upon the power of statistical tests. Attention was focused on the F test of the single factor analysis of variance. Formulas were derived to show the relationship between the noncentrality parameters for analyses using true scores and those using observed scores. The effect of the reliability of the measurement and the sample size were thus demonstrated. The assumptions of classical test theory were used to develop formulas relating test length to the noncentrality parameters.

The cost of an experiment was analyzed in terms of a fixed cost per subject and a variable cost dependent upon test length. Computer programs were written to estimate the power under all permissible allocations of resources to sample size and test length. The program results indicate which of the permissible allocations will result in maximum power.

◆ ◆ ◆

Mean Differences and Association

GEORGE W. JOE and HARRY E. ANDERSON, JR., University of Georgia

Recently, several articles have emphasized the desirability of reporting "converted correlation coefficients" derived from "tests of significance of mean differences" as an aid to the interpretation of these significance tests. The first part of this paper is an extension of this idea to Multiple Correlation, R, and the Mahalanobis D^2.

The second part of this paper presents a derivation of transformation of the weights, a, of a discriminant function for g-groups into a set of beta weights, β, in which pseudo-variates are not assumed.

The Effects of Using Fallible Variables in the Analysis of Covariance

ANDREW C. PORTER, Michigan State University (Robert L. Ebel)

The conventional analysis of covariance is not an appropriate method of analysis when the covariable contains errors of measurement. The inappropriateness rests upon the bias of the least-squares estimate of the common slope of treatment group regression lines defined by the true parts of the variables. The present paper is concerned with the properties of the sampling distribution of the test statistic arising from a modification of the analysis of covariance procedure proposed by the writer. The modification is computationally consistent with the analysis of covariance; however, the covariable is estimated true scores rather than the observed fallible scores.

When the population reliability of a fallible covariable is used for calculating estimated true scores, the estimated true scores provide a covariable which exactly corrects for the bias of the least-squares estimate of the structural relation. The traditional model for the analysis of covariance holds for use of such estimated true scores and the distribution of the F test statistic follows the theoretical F distribution. When the reliability of the fallible covariable must be estimated, an additional source of variation is introduced into the least-squares estimate of the structural relation.

A Monte Carlo approach was used to generate sampling distributions of the test statistic arising from the modified analysis of covariance. The generated distributions were compared with the corresponding theoretical F distributions to determine whether the theoretical F served as a useful reference distribution.

The investigation of the use of estimated true scores as a covariable, when they are based on an estimate of reliability, indicated that the additional source of variation was not sufficient to defeat the practical usefulness of the procedure. The data clearly recommended that estimated true scores should be used as a substitute covariable in the analysis of covariance where the observed covariable has a population reliability between the values of .5 and .9.

◆ ◆ ◆

An Empirical Investigation of the Combined Effect of Heterogeneity of Variance and the Existence of Block-Treatment Interaction on the F-Test Under Permutation for the Randomized Block Design

LARRY E. TOOTHAKER, University of Wisconsin (Frank B. Baker)

The purpose of this research was to investigate the combined effect of heterogeneity of block-variance and the existence of block-treatment interaction on the size and the power of the F-test under permutation in the randomized block design. The study was begun by obtaining the empirical distributions of variance ratios under permutation by means of a Monte Carlo program. The permutation distributions were compared with the normal theory distribution by determining the proportion of F-ratios greater than the upper α-point under normal theory for two randomized block designs with one observation per cell. Estimates of the size were obtained for various specified levels of block-variance heterogeneity and block-treatment interaction for various values of normal theory size. Estimates of the power were obtained for the same levels of the violations of assumptions as were used for the size for various levels of normal theory power and $\alpha = .05$. The empirical results for the size and power of F_t under permutation showed a curious increasing and decreasing or 'humping'. For non-null levels of each violation of assumptions, the humping occurred over increasing non-nullity of the other violation. Further investigation of the problem showed that the increasing block-treatment interaction was changing the variance heterogeneity in that it was changing the average over all possible permutations of a measure of variance heterogeneity. When the nominal values of size were altered accordingly, closer agreement with the empirical size was obtained; however, the corrected nominal values did not have the humping of the empirical size. Most empirical evidence points to a combined effect of the violations studied. The most dominant factor affecting the empirical size and power would seem to be the changing of variance heterogeneity by the increasing block-treatment interaction.

11:00 to 11:55 A.M. **SESSION 16.3** **DIVISION A**

SCHOOL BOARD-SUPERINTENDENT RELATIONSHIP

CHAIRMAN:

Edward S. Hickcox, Ontario Institute for Studies in Education

PAPERS:

A Cross-Cultural Study of Work Relationships

R. OLIVER GIBSON, State University of New York at Buffalo

This study reports on an analysis of materials in the Human Relations Area Files supported by the Bureau of Research of the U. S. Office of Education. The purposes are to identify materials relevant to selected aspects of employment in formal social systems, to analyze them in terms of general systems analysis based upon such concepts as 'social exchange' and 'contract' and, as a result, to reformulate and refine the conceptualization.

In recent years 'social exchange' has been developed in the literature as a useful analytic concept for dealing with the interaction of persons in social systems. A related concept that proves useful in dealing with exchange is that of 'contract.' Presumably the nature of the exchange and the form of the contract would vary from social system to social system along with differences in the general norms of the systems, the nature of status ascribed to the work position, the formalism—informalism of relationships, and the like. Benedict's "Apollonian—Dionysiac duality" offers a possible dimension of variation.

The Human Relations Area Files have been searched under such headings as "Exchange", "Labor", "Mutual Aid", and "Contracts." The materials have been coded and analyzed in terms of the formulation of Gibson in *A General Systems Formulation of Work Relationships.* Exploratory analysis has been in terms of the following social systems: United States, Canada, England, Zuni, and Kwakiutl.

The present study is seen as throwing light upon personal relations in formal organizations as a form of social system, and as having relevance for occupational planning and education.

On the basis of the above analysis, the initial formulation is further developed, several more specific hypotheses are proposed, and implications for personnel administration and occupational planning explored briefly.

◆ ◆ ◆

Personality Characteristics of School Superintendents in Relation to Their Willingness to Accept Innovation in Education

HOMER M. JOHNSON, GEORGE M. CARNIE, and CLIFFORD J. LAWRENCE, Utah State University

On the assumption that the superintendent of schools must play a key role in meeting the challenge of change through implementation of innovation in education, we felt a need to seek better ways of identifying innovators. A review of research indicated that people in various roles can be distinguished by unique personality profiles. Our problem, then, was to determine whether or not superintendents who implement change could be identified by the uniqueness of their personality traits. The purpose of our research was to determine whether a correlation existed between personality characteristics of school superintendents and their willingness to implement innovations in education. The research also tested for differences in personality factors between high innovative and low innovative superintendents.

Data were gathered from 93 Idaho superintendents and 71 superintendents, identified as innovators, from 12 other states. The superintendents responded to Cattell's 16 P. F. Questionnaire and a three part scale on innovation.

A significant multiple correlation of .52 was found between the 16 personality factors and the composite innovation score. Six personality factors (C, F, H, I, L and Q_4) made a significant contribution to the differences in the composite score. A significant multiple correlation of .52 was also found between the 16 personality factors and the education practice score with factors B, E, F, H, I, N and Q_3 making a significant contribution to the differences in the practice scores. We also found that high innovative and low innovative superintendents differ significantly on six personality factors: A, E, H, M, Q_1 AND Q_4.

1. There is a significant correlation between personality characteristics of superintendents and their willingness to accept and implement change in education.

2. There is a significant difference between the personality characteristics of the high innovative superintendents and the low innovative superintendents. The high innovative superintendents are more outgoing, more assertive, more venturesome, more imaginative, more experimenting, and more relaxed than the low innovative superintendents.

3. It seems likely, then, that personality characteristics, if adequately determined, can be useful in identifying and predicting innovative administrators in education.

◆ ◆ ◆

The Effect of Five Variables on Educational Decisions as Perceived by Board Members and Local School Superintendents

JAMES B. KENNEY and HASKIN R. POUNDS, University of Georgia

Five questions were verbally posed to 30 local superintendents of public eduation and to 92 members of their boards of education in 3 southeastern states in an effort to determine if there was a consensus of perception between the two groups relative to the effect of five variables on public education at the community level.

The distribution of *Ss* was as follows: State #1, 10 superintendents and 31 board members; State #2, 10 superintendents and 30 board members; and State #3, 10 superintendents and 31 board members. All school systems included were medium or small as defined by pupil ADA.

Questions posed to superintendents were as follows:

I. On the basis of your knowledge of the local community, what "social rank" (as defined by Feldvebel) do you perceive each of your board members as occupying?

2. On the basis of your knowledge of the local community, what "economic rank" (as defined by Galbraith) do you perceive each of your board members as occupying?

3. On the basis of your knowledge of the local community, would you classify the community as "liberal or conservative" (as defined by Gouldner)?

4. Are decisions affecting local education often made in a "social setting" by members of your board of education? (As opposed to decisions made in regularly constituted board meetings).

5. Is there a strong "informal power group" (as defined by Kinbrough) in your community that influences local educational decisions to a significant degree?

Likert type scales were constructed and board members were asked to respond to the same questions appropriately rephrased with the following results.

1) Superintendents perceived members of their board of education as occupying significantly higher ($p < .01$) positions of social rank in the local community than did the board members when viewing the position of their superintendent.

2) Superintendents perceived members of their boards of education as occupying a significantly higher ($p < .001$) position on the economic scale of the local community than did board members when viewing the position of the superintendent.

3) There was no statistical difference in perception of board members and superintendents relative to the "liberal-conservative" classification of their community with virtually equal distribution of responses in each category.

4) Superintendents and board members did not differ significantly in their responses to educational decisions being made in social settings versus formal board meetings.

5) Superintendents perceived the influence on educational decisions of

an informal power group in their community as being significantly stronger ($p. < .01$) than did their board members.

Substantiating arguments drawing on Wellman, Gaurley, Levin, Dahl, Rossi, and Fisher are presented in the discussion. Particular attention is given to the apparently contradictory responses evidenced by Questions 4 and 5 above.

◆ ◆ ◆

Variables Influencing the Vulnerability of Superintendents and Their Actions and Policies

DAVID J. MAHAN, St. Louis Public Schools

This historical case study of the scientific management era focuses (1) on the relationship between the vulnerability of urban superintendents and their adoption of financial economy, and (2) on variables determining the state of vulnerability. Research views actions and policies of superintendents in relation to economy and cost accounting principles advocated by proponents of scientific management and in relation to Callahan's vulnerability thesis.

Superintendents of the St. Louis public schools during 1910-1920 were subjects. Numerous conditions prior to 1910 in St. Louis tended to predict with a high probability the adoption of financial efficiency principles. Data were collected from several primary sources including board proceedings, annual reports, national journals, and daily newspapers.

Analyses of data indicate that contrary to prediction, superintendents did not focus their attention on financial economy and assume the role of a business executive. Instead they took exactly opposite positions. This rejection was associated with a low state of vulnerability. Several factors—legal, community and personal—were responsible for this lack of vulnerability. A legal charter reserved specific administrative powers for the superintendent and provided for a four year term during which salary could not be reduced. The community did not subordinate the maintenance of educational quality to financial economy. Civic leadership and an intelligent press crusaded for added expenditures to improve the quality of the school system. Personal factors included national prominence, close association with community power figures, and offers of another prestigious job while superintendent.

The study demonstrated that in the absence of conditions of vulnerability, superintendents did not automatically adopt financial principles urged by a national movement. Legal, community, and personal factors reduced vulnerability. One major finding of the study is the suggestion that personal variables of incumbents may be a major factor in determining the vulnerability of the superintendency. Previous investigations have focused mainly on legal powers and community pressures.

11:00 to 11:55 A.M. SESSION 16.4 DIVISION D

JOINT SESSION WITH NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION
TEST THEORY

CHAIRMAN:

Arvo E. Juola, Michigan State University

PAPERS:

Expectable Reliability in Objective Tests

ROBERT L. EBEL, Michigan State University

This paper presents a theoretical analysis of variations in the expected reliability of tests of standard length composed of items offering varying numbers of alternatives. One conclusion reached is that a test composed of four choice items is likely to be only slightly more reliable than a three choice item.

The analysis rests on the assumptions (1) that the ideal mean score for an objective test is located midway between the maximum possible score and the expected chance score, and (2) that the reasonably expectable standard deviation for such a test is one sixth of the range between the maximum possible score and the expected chance score. On these assumptions, and given the number of items in the test and the number of choices per items, KR-21 reliability estimates can be calculated. These are the quantities referred to as expectable reliability coefficients.

One of the inferences from this study is that there is a definite limit on the expectable reliability of a test. Improvement in item quality is not likely to raise the reliability above this limit, though poor item quality will lead to reliabilities well below the limit.

A formula relating KR-21 to n, the number of choices per item and k, the number of items in the test has been developed and will be presented.

◆ ◆ ◆

A Monte Carlo and a Logical Analysis of Multitrait-Multimethod Validational Processes

MILTON D. JACOBSON and JOHN J. YOUNG, University of Virginia

Personality-trait tests are widely used and are being produced in abundance. The question then arises, "Which tests are good or valid?"

Campbell and Fiske advocated the use of multitrait-multimethod intercorrelation matrices and developed four criteria which provide for both convergent and discriminant validation of psychological traits.

Logical and statistical (probabilistic) difficulties were recognized in multitrait-multimethod validation. It was the purpose of this research to investigate these two difficulties, and these have been treated separately.

Part I of this study investigated the appropriateness of using F statistics developed for intercorrelation matrices to validate data obtained from small sample sizes. Monte Carlo procedures were used to generate empirical distributions of these statistics to evaluate their robustness under non-null conditions of method bias.

In Part II, a logical analysis was made of Campbell and Fiske's four criteria which provided for: convergent validity, discriminant validity, trait-method units and multitrait-multimethod measurement. This analysis examined the common sense notion as well as the technical concepts of *test, validity* and *reliability* and where possible transformed the results into symbolic logic to make the conceptual properties and relations as clear as possible to determine if these criteria are contingently or logically necessary.

These F statistics were not robust but were sensitive to non-null contributions of method and method-trait bias. The combinations of non-null conditions which produce minimal distortion of the statistics were detrmined, Tables of these conditions, which prescribe the usefulness of multitrait-multimethod F statistics for small sample sizes are presented.

It was concluded that only criterion 1 (convergent validity) seems to be logically necessary. Criteria 2-4 (discriminant validity) do not seem to be based on a priori grounds and must be justified as utilitarian and contingent but not necessary. As such, criteria 2-4 are severely weakened.

◆ ◆ ◆

The Principal Components of Multiple-Guttman Scales

JON MAGOON and RICHARD C. COX, University of Pittsburgh

The construction, interpretation, and evaluation of Multiple-Guttman Scaled measuring instruments suggests a need for reference to the principal component structure of this type of data. This empirical study represents an attempt to provide a description of the principal component structure of Multiple-Guttman Scaled data when the important attributes of these data are known.

Multiple-Guttman Scaled data has three important attributes; (1) the number of independent scales within the data, (2) the reproducibility of the multiple scales, and (3) the distribution of responses to scaled items within a scale or redundancy. A pattern-generating computer program was written to produce item response data for various combinations of multiplicity of scales, reproducibility, and redundancy. The size of the independent scales and the number of scales in the data varied from one, to two, three and five. The reproducibility, simulated by introducing random error into perfect scale patterns, varied from .80 to 1.00. The redundancy of responses within each scale took on U-shaped, flat, round and leptokurdic distributions. The distribution of responses was expected to exert a most important effect on the principal component structure, since items of middle difficulty have the highest relative variance.

For each combination of the multiple scale parameters, the items constituting the multiple scale pattern were intercorrelated. When the inter-item coefficient is a tetrachoric, the items of single scales are highly related as a rule. The matrix of tetrachoric correlations is just as often non-Gramian however, and thus difficult to interpret from a principal components viewpoint. Usage of phi intercorrelation coefficient avoids the problem of a non-Gramian matrix.

The analysis of the principal component structure of eighty different item intercorrelation matrices derived from Multiple-Guttman Scaled data suggests the structure which can be expected under a wide variety of specified conditions. These findings provide a better understanding of the relations which exist in Multiple-Guttman Scale data, and point out several relations which exist between Multiple-Guttman Scales and the principal components of the same data.

◆ ◆ ◆

The Approximate Sampling Distribution of Kuder-Richardson Twenty

DOUGLAS A. PENFIELD, University of California, Berkeley

In 1941, Cyril Hoyt demonstrated that an internal consistency measure of reliability could be established using Analysis of Variance procedures. This was shown to be equivalent to the reliability coefficient (K-R 20) derived by Kuder and Richardson in 1937. From the derivation of the sampling distribution of K-R 20 using a random effects model of the Analysis of Variance, Leonard S. Feldt established the probability statement, $P(1 - (1 - \rho_{20})/F_L \leq r_{20} \leq 1 - (1 - \rho_{20})/F_U) = 1 - \alpha$.

The sampling distribution of K-R 20 for five distributions is investigated by the author for specified values of ρ_{20}. Using Type 12 sampling described by Lord, a 7094 IBM computer was utilized for drawing subjects and items at random from a known population. For each subject (N) x item (n) matrix drawn, K-R 20 is computed. The sampling distribution of K-R 20 is investigated for N and n taking on all combinations of the sample values 5, 10, 30 and 50. The empirical distribution of K-R 20 is then compared at specified percentile points with the corresponding percentile points of the theoretical distribution derived by using Feldt's probability statement mentioned earlier.

Once the theoretical and empirical distributions are established, the Kolmogorov-Smirnov Goodness of Fit test is used to determine whether the two distributions resemble each other.

For the populations investigated in this study, the fit between the empirical and theoretical distributions is very poor. As the number of items in the sample increases, the fit tends to improve between the expected and observed reliability distributions. A discussion of the empirical sampling distributions of K-R 20 for each of the derived populations is presented.

11:00 to 11:55 A.M. SESSION 16.6 DIVISION E

GROWTH RELATIONSHIPS AND PUPIL DEVELOPMENT

CHAIRMAN:
Thomas P. Nally, University of Rhode Island

PAPERS:

Some Correlates of Reading Readiness among Children of Varying Backgrounds

BARBARA H. LONG, Goucher College and EDMUND H. HENDERSON, University of Delaware

The purpose of this study was to investigate the correlates of reading readiness among children of varying background. Subjects were 192 beginning first-graders selected on the basis of sex and race from 13 schools in a rural Southern area. An attempt was made to control for socioeconomic level by use of the chief earner's occupation (half of each group were in classes 6 and 7 on Hollingshead's Occupational Scale, half in classes 1 through 5). Subjects were administered the Children's Self-Social Constructs Test (a non-verbal instrument providing 12 measures of self-concept; split-half reliabilities from .48 to .75, median .73 among 100 first-graders), and the Metropolitan Readiness Test. Teachers rated them on 24 classroom behaviors (summed to provide a single score; split-half reliability, .95). Information about age, family structure, and kindergarten/Head Start attendance was collected.

Intercorrelations for the total sample revealed 8 of the 17 variables to be significantly related to readiness scores ($p = .05$ or better). The three best predictors (multiple $r = .62$) were kindergarten attendance, classroom behavior, and chronological age. Also related to greater readiness were three self-concept variables: (a) greater realism of self-concept as to size; (b) greater preference for mother; (c) greater self-teacher distance, and two family variables: (a) father presence and (b) fewer siblings.

Since the groups differed significantly on a number of these variables, separate correlations for the Negroes and whites and boys and girls were calculated. Family variables were not related to readiness with race and kindergarten attendance controlled. For the three self-concept measures, differential relationships were found among the four groups. In each group, kindergarten attendance, classroom behavior, and chronological age were most predictive (multiple r's ranged from .51 to .71). The results thus support the importance of kindergarten experience, and suggest the relevance of teacher judgments for reading readiness.

Dimensions of Creativity in Elementary School Children

JAMES H. BEAIRD, Teaching Research Division of the Oregon State System of Higher Education (T. Antoinette Ryan, Oregon State University)

Previous Creativity research has awakened the educator to the realization that school success is not totally dependent upon scholastic aptitude alone. Getzels and Jackson initially reported findings that highly creative students often performed as well as their more intelligent cohorts in many achievement situations, an observation verified by Torrance's research. Recently, several writers have published texts which outline teaching strategies designed to enhance and build upon the creativity of school children.

While the soundness of these teaching strategies and the research base upon which they are predicated cannot be questioned, there is a growing body of evidence which suggests that such strategies must be augmented with closer attention to certain noncognitive characteristics of learners.

The purposes of this study was to (1) identify the dimensions of creativity exhibited in three classes of objects (written, art manipulatable) produced by elementary school children; and (2) to relate these dimensions to cognitive and noncognitive characteristics of the producers.

Sixth grade students (n = 25) enrolled in a single self-contained classroom, each produced a written story entitled "A Walk in the Woods," a diorama illustrating the story, and a usable object constructed from a kit of provided materials. Judges representing the areas of psychology, education, music, art, and the humanities rated the relative creativity exhibited in each sample of objects following multidimensional scaling procedures described by Torgerson (1958).

Two types of measures of student characteristics were utilized. Personality characteristics were assessed with the Children's Personality Questionnaire. A battery of tests designed to measure cognitive skills was selected from the Educational Testing Service Reference Kit for Tests of Cognitive Factors and administered.

Three salient dimensions were interpreted for each class of student products. Interpretations of each dimension required attention to both cognitive and personality student characteristics. Not one dimension could be labeled as being a purely cognitive factor.

It was apparent that individuals who produce creative stories possess a set of characteristics unique from those who produce creative art or creative kit objects. This suggests refinement of school testing programs, especially in the utilization of personality measures, a heretofore sensitive area in public education. Secondly, the presence of personality correlates to factors of creativity suggests that present strategies for enhancement of creative behavior may require modification.

An Analysis of Developmental Changes in the Self-Concept of Negro and White Children Grades 3-12.

PEARLINE PETERS YEATTS, University of Georgia (T. Antoinette Ryan, Oregon State University)

The major purpose of this study was to examine the different factors of self-concept which emerge in relation to sex, race, and socio-economic status at different age levels from grades three through twelve.

An additional purpose was to create normative data which could be used in assessing the self-concept, based upon the dimensions of self which emerged for the subjects used in this study.

Gordon's "How I See Myself" self-report scale was administered to the student population, grades three through twelve, of a cosmopolitan North Central Florida public school system. The total population included 8,979 subjects. Approximately one third of the population was Negro, and all socio-economic levels were represented.

Responses to the self-report instrument with the variables of grade, race, academic achievement, and socio-economic status were, with the aid of an IBM 709 computer, intercorrelated. The demographic data were added prior to making further investigations. A principal-components analysis was made, using units in the diagonals. The factors were then rotated by means of varimax rotation.

The scores were then summed for the purpose of acquiring the means and standard deviations of each factor for each group where the factor emerged.

The factors emerging supported the postulate that the self-concept is not a unitary-concept, and that the conceptions one holds in regard to self vary with age and sex.

Each group was unique in regard to the dimensions of self emerging, yet common factors emerged for each group indicating common concerns in regard to self.

The normative data obtained provide a base against which, for the first time, the self-concept of children of various ages, sex, and race can be compared.

◆ ◆ ◆

Forecasting Academic Status

RICHARD WOLF, University of Southern California

A great deal of information is collected about each student as he progresses through school. This information, in the form of test scores, teacher's marks, and personal and social data, is usually stored in cumulative folders. The collections and storage of such information is usually highly systematic. Unfortunately, there is little in the way of systematic use of such information. The present study represents one attempt to treat such data and prepare it for systematic use. Third grade and sixth grade standardized test scores in reading vocabulary, reading comprehension, arithmetic fundamentals, and arithmetic reasoning were gathered for over 700 students in several school districts representing a wide socioeconomic distribution. Third grade intelligence test scores were also obtained. These data were analyzed through the use of a stepwise multiple regression procedure to obtain the best set of weights for using third grade test data to forecast sixth grade status in each of the four subject areas. The resulting multiple correlations ranged from $+.65$ to $+.71$ and held up under cross-validation.

These results indicate that third grade test data can be used to forecast sixth grade achievement with a fairly high degree of success. A manual has been developed and tested with third grade teachers. Each teacher is furnished with simple forms and instructions on how to forecast the sixth grade achievement for each of her students. These forecasts are estimates of the level of achievement in each of four subject areas in the sixth grade under the condition that no special treatment is accorded the student. Teachers using the forecasting procedures are able to identify students headed for failure and intervene to destroy such a forecast. Through early identification of potential failures, it is hoped that such forecasts can be destroyed through special instructional efforts. The likelihood for success of such efforts is high since there is a three year period in which teachers have time to work with students who are forecasted to fail.

12:00 M to 12:55 P.M. SESSION 17.4 DIVISION C

TEACHING THE DISADVANTAGED II

CHAIRMAN:

Milly Cowles, University of South Carolina

PAPERS:

A Comparison of Learning Problems and Responsiveness to Remediation in Culturally Disadvantaged and Advantaged Youngsters

SEYMOUR FESHBACH, University of California, Los Angeles
(Norma D. Feshbach)

The purpose of the present study was to evaluate a remedial program, and to determine whether the learning problems presented by a culturally disadvantaged population were fundamentally different from the learning problems presented by a sample of middle class children who constitute the basic source of most of our information about learning disorders.

To obtain evidence bearing on this issue, 20 elementary age and 20 junior high school age children from culturally disadvantaged areas, and who manifested significant retardation in some basic skill, were bussed to the Psychology Clinic School on the U.C.L.A. campus for a six-week remedial program conducted during the summers of 1966 and 1967. The children attended classes with advantaged youngsters with learning problems. A number of measures of achievement in basic academic skills, intelligence, cognitive and perceptual functions, motivation, and attitudes were administered before, during, and after the remedial program. As part of a related study, 60 disadvantaged children were grouped into matched triplets; one child from each triplet was bussed to the Clinic School for the academic year, a second child was randomly assigned to an Educational Enrichment program conducted in his own home school, and the third child to a Control group. The test battery was also administered to these children and to a matched group of advantaged youngsters attending the Clinic School.

Some of the major findings indicated that:

1. advantaged and disadvantaged youngsters placed similar values on academic attainment, but the disadvantaged child experienced greater anxiety regarding school-related matters, was poorer in auditory discrimination, and was more field-dependent;

2. both advantaged and disadvantaged children attending the Clinic School made significantly greater gains in Achievement than did the other groups. Simply placing the disadvantaged child in a more congenial environment resulted in higher achievement scores, less anxiety and greater vocational aspirations on the initial testing.

Development of an Instructional System

RALPH J. MELARAGNO and GERALD NEWMARK, System Development Corporation

An instructional system was developed empirically to teach reading-readiness concepts to first-grade Mexican-American students. The objective of the study was to develop operational procedures for improving classroom instruction; specifically, to investigate the extent to which procedures of "shaping" materials for individuals (programmed instruction) could be applied successfully to a total classroom configuration in which interactions among students, materials and teacher are considered.

With teachers as an integral part of the research team, classroom variables were subjected to the evaluation-revision strategy until an instructional system was developed which achieved specific objectives with 90% of learners. The study involved three phases.

A. Identification through classroom observations and testing of a reading problem for Mexican-Americans: understanding the concepts conveyed by 10 words used frequently in reading instruction (top, bottom, over, under, first, middle, last, alike, different, and underline).

B. Empirical development of seven different instructional procedures: (1) *Teacher-led Orientation* of an entire class; (2) *Teacher as Tutor* of four or fewer learners; (3) *Classroom Small-Group Instruction* with learners paired to help one another while using booklets and tape recordings; (4) *Structured Individual Tutoring* with upper elementary tutors trained to use booklets and tape recordings; (5) *Unstructured Individual Tutoring* with upper elementary tutors choosing methods and materials; (6) *Small-Group Tutoring* with upper elementary tutors helping two or three learners; and (7) *Parents as Tutors* of their own children.

C. Empirical development of the total instructional system, integrating instructional and support activities. The experimenters, aided by the teachers, conducted the first trials; participating teachers were assisted by the experimenters with subsequent trials; and naïve teachers performed the final trials unassisted.

Results showed the system to be highly effective, but its successful utilization is dependent on supplementing school resources. A number of future research studies are indicated. The evaluation-revision strategy was validated for improving classroom instruction.

◆ ◆ ◆

The Effect of Kindergarten Experience on Reading Achievement of Disadvantaged Urban Negro Children

COLEMAN MORRISON and ALBERT J. HARRIS, The City University of New York

This paper deals with the effect of kindergarten vs. non-kindergarten experience on achievement when four methods of reading instruction are used. It is part of a larger project designed to compare the effectiveness of reading instruction provided disadvantaged urban children in the New York public schools.

Kindergarten experience in this study is defined as attendance in kindergarten for a period of 101 to 200 half-days.

Achievement tests included Stanford Primary I Battery for grade one, and the Metropolitan Primary II for grade two. The methods of reading instruction included 1) basal reading; 2) basal reading plus phonovisual; 3) language experience; and 4) language experience plus audio-visual materials.

Children participating in this study began their kindergarten experience in the fall of 1963. The kindergarten curriculum in the schools attended by these children should be considered "unstructured" and typical of kindergarten programs in a majority of schools throughout the country. Approximately two-thirds of the children had kindergarten experience.

For all methods combined, children having kindergarten experience achieved significantly higher scores on three reading subtests of the Metropolitan Achievement Test at the end of second grade than did children without such experience. There was a significant relationship between kindergarten experience and the Language Experience approach on all sub-tests of the Metropolitan, and a significant relationship between kindergarten experience and the Language-Experience-Audio-Visual method on the reading sub-test.

In the Basal Reader Method the kindergarten group scored slightly but not significantly higher. In the Phonovisual Method none of the three differences was significant; two were in favor of the kindergarten group, and one (Reading) favored the non-kindergarten group.

Kindergarten experience was used with four reading readiness measures in a multiple correlation with each of the outcome measures. Kindergarten experience contributed substantially to the multiple repression equation for oral vocabulary, but had zero weight in the regression equations for the reading subtests.

It may be concluded that the kind of kindergarten experience which these children had, influenced their vocabulary development somewhat. Children with kindergarten experience had a slight advantage in the language-experience methods, but did not show a measurable advantage in the learning of specific reading skills in skills-centered methods.

An Evaluative Study of Prekindergarten Programs for Educationally Disadvantaged Children: Follow-Up and Replication

LOUIS T. Di LORENZO and RUTH SALTER*, New York State Education Department

This will be the second report to AERA on a longitudinal evaluation of Prekindergarten Programs for Educationally Disadvantaged Children; the first was given at the 1967 Annual Meeting.

The initial report detailed findings on changes in capacity to learn, as measured by the Stanford-Binet Intelligence Scale, and in receptive vocabulary, as measured by the Peabody Picture Vocabulary Test, for 570 experimental and control children. The tests were administered on a pre-post basis in 1965-66, the first year of the four-year project.

The second-year report will include the following: (1) An analysis of scores on Metropolitan Readiness Tests administered to the first-year sample at the end of kindergarten; (2) an analysis of I.Q. and language scores for the 650 prekindergarten children in the second year of the project; and (3) a comparison of the findings for the two years. Factors to be considered in the several analyses are sex, race, socioeconomic status, and type of program.

The prekindergarten programs, located in eight school districts, represent divergent points of view in preschool education. Several are traditional nursery school programs, one employs machine teaching, and another is using an adaptation of the language pattern drills developed by Bereiter and Englemann. Two of the programs have heterogeneous classes of disadvantaged and non-disadvantaged children; the others have homogeneous classes. All programs have as their goal the improvement of capacity to learn, social development, self-concept, physical development, and attitudes toward school.

12:00 M to 12:55 P.M. SESSION 17.6 DIVISION A

VARIABLES RELATED TO EDUCATIONAL OUTCOMES

CHAIRMAN:

Richard Morrow, University of Wisconsin

PAPERS:

Teacher Attitudes Toward the Local School System: Development of the TORS

R. ROBERT RENTZ and JAMES B. KENNEY, University of Georgia; JEFFREY L. LORENTZ, Bibb County, Georgia, Board of Education

Educational innovations instituted by the school system's central office must generally be implemented by the classroom teacher, and innovative classroom procedures devised by the teacher (depending on their scope) ultimately require central office approval. Clearly, the attitudes and opinions of the classroom teacher with respect to system-wide central office policy, procedures, and personnel are critical to the success of these innovations. The *Teacher Opinion Rating Scale* (TORS) was developed to assess the attitudes of teachers toward the various aspects of the school system and central office, as an important part of the evaluation of Title I activities in Bibb County, Georgia.

A committee of teachers, principals, administrators, and central office personnel initiated a pool of items considered pertinent to the foregoing objectives. After subsequent deletions, revisions, and additions by the investigators, a set of 85 items was generated under the following five categories: (1) attitudes regarding materials; (2) attitudes regarding facilities; (3) organization within the school; (4) coordination and communication; (5) administration. The items required respondents to agree or disagree by selecting the appropriate position along a five-point scale (1 = strongly disagree; 5 = strongly agree).

This initial form of the TORS was administered to a randomly selected sample of 136 teachers from the entire school system. These item responses were factor analyzed and rotated to the varimax criterion. Six rotated factors were given preliminary interpretations and nine items were eliminated because of low communalities. The remaining 76 items constituted *Form B* to be used in a system-wide administration.

Subsequently, *TORS-B* was sent to every teacher in the system; 1329 were completed and returned, representing 89.9 percent of the system's teachers. These data were factor analyzed using two rotation procedures: Varimax (orthogonal) and Maxplane (oblique). The oblique factor structure more closely approximated simple structure. Four large factors were considered important and interpretable. These were identified as: (I), Extra-School Coordination and Communication; (II), Criticism of School Conditions; (III), Non-Classroom School Satisfaction; (IV), Educational Orientation. The 47 items best defining these four factors were considered of sufficient generalizability for use in other school systems, and possibilities for use in system-wide evaluations were discussed.

◆ ◆ ◆

Economy of Scale in the Production of Selected Educational Outcomes

MARVIN C. ALKIN, U.C.L.A., CHARLES S. BENSON, University of California, Berkeley, and ROBERT GUSTAFSON, California State Board of Equalization

In this study, the authors sought to examine the concept of economy of scale as it relates to school systems. While there have been at least two studies related to economy of scale for education, (Nels Hanson, 1964; Werner Hirsch, 1959) neither had adequate controls for the educational outcomes being produced. That is, in one of the studies noted, the question asked was most nearly: "How is size of district related to how much school districts spend per pupil?" (Hanson). In the other study, the criterion measures were so much a direct function of student inputs and system characteristics as to make them inappropriate for consideration as outcome measures (Hirsch). We sought in this study, to examine the relationship between size of school district (ADA) and several measures of educational outcome (mean and median reading and arithmetic scores at the 5th grade), while controlling statistically for student and financial inputs (socio-economic characteristics of communities and dollar expenditure per ADA, respectively).

The population included all California school districts with elementary school programs and which were within Standard Metropolitan Statistical Areas of the State in 1960. A grouping variable was used to distribute the 255 school districts by type of district. Elementary District 1-6 (GS 1), Elementary District K-6 (GS 2), Unified District (GS 3). Separate analyses were made for the three with the following number of cases in each grade span: GS 1: 34 school districts, GS 2: 176 school districts, GS 3: 45 school districts.

Within each of the three grade spans, stepwise multiple regression equations were used to predict each of the 4 criterion measures from 8 socio-economic attributes of the census area of the district, and from a measure of the per pupil financial input to the district. Then, for each of the 12 equations (3 grade spans, 4 criterion measures) we "forced" an additional independent variable—size of district (ADA)—to determine whether the introduction of the new variable increased the multiple R of the equation. And, if predictability was increased, we hoped to examine the nature of the linear relationship as expressed by the regression coefficient for the forced variable.

Multiple R's of the 12 original equations ranged from .592 to .857. Introduction of the size variable yielded the following tentative conclusion: Given equivalent financial resources and comparable socio-economic characteristics, there is no evidence that greater educational benefits are produced in larger school districts with elementary school programs (within the size range of districts included in this study).

◆ ◆ ◆

Personality Orientation and Administrator Behavior

D. L. DUNCANSON, University of Idaho

Relating personality or non-decision-making behavior with administrator decision-making behavior has intrigued numerous scholars for a long time. The developing of measures that allow prediction of success has lured many to investigate the area. The results of this study do not warrant the use of either of these two measures of personality for purposes of selecting and/or retaining aspiring administrators in preparation programs. There was no established connection between administrator personality and administrator decision-making behavior.

The in-basket technique can be advantageously used to determine patterns of administrator decision-making. Factor analysis can be used to group discrete behaviors into patterns.

The results of this study are very similar to portions of the Hemphill, *et al.* study. The first factor in both studies was related to the exchanging of information. This represents a major aspect of the administrator's behavior. It suggests that facility in the various skills required to exchange information is necessary to the administrator.

Secondly, both studies point out the importance of the continuum of decision-making vs preparation for decision-making. This was one of Hemphill, *et al's* secondary factors that cut across their whole study. In a similar manner, this was an important aspect of factors #1, 2, and 4 in the current study. The administrator must, then, also be competent in the area of making decisions.

The current study found at least one contrasting phenomenon that was not present in the Hemphill, *et al.* study. There was a focus on whom the administrator did *not* involve. The results show strong negative loadings on tests of working with certain groups, i.e., factor 2 with subordinates, factor 3 with students, and factor 4 with outsiders. It is also observed that the administrator *made decisions* without involving subordinates (in factor 2) or outsiders (in factor 4); while in factor 3, he excluded the students but did not necessarily make any decisions nor take any terminal action.

12:00 M to 12:55 P.M. SESSION 17.7 DIVISION E

COLLEGE STUDENTS: VALUES AND ATTITUDES

CHAIRMAN:

John R. Braun, University of Bridgeport

PAPERS:

College and Student Values: A Longitudinal Study

WALTER L. THOMAS, Project on Student Values (T. Antoinette Ryan, Oregon State University)

The study attempted to provide information on the following questions:

A. Do church-related, private non-church, and public colleges and universities differ in the types of students they admit, drop, and graduate in terms of personal values?

B. Are values related to college grades?

C. Do the above types of institutions see changes in student values during four years of college?

D. Are "leavers" and "stay-ins" distinguishable in terms of values?

Students numbering 4,595 were sampled on a stratified and proportionate basis among 16 colleges and universities. The *Differential Value Profile* was administered to freshmen in 1963, and later as seniors in 1967. "Leavers" were compared with "Stay-ins" on the basis of their freshman scores. "Stay-ins" were studied for value change by comparing their DVP scores four years later. Inter-institutional and male-female comparisons were made on each major effect in the study. Value indices were also used as freshman grade predictors. Analysis of variance and multiple regression analysis were used respectively to test 323 null hypotheses.

Of 288 null hypotheses of no value difference, 122 were rejected at the .05 level. Of 35 null hypotheses of no relationship (values and grades), 14 were rejected at the .05 level.

Values can be used to early-identify potential college drop-outs. A drop-out value syndrome is identifiable relative to a given campus. Colleges do see value change, but not always in the directions purported. Design problems in the Jacob Studies are corrected, and the Jacob conclusions refuted. Values are useful predictors of college grades, particularly in augmenting ACT and converted H. S. Rank predictors. A student-to-college "fit" can be significantly augmented with value indices on both student and college. Small colleges see more value change than universities. Women have more value change than men. Value change is more related to *size* than *type of control* of an institution.

The Relationship of Personality Factors to Scientific-Mathematical Aptitude

CAROLYN NEAL HEDLEY, University of California, Santa Barbara

University of Illinois freshmen (N-347) entering the School of Education were administered a series of tests, the MMPI, the Kuder Preference Record measuring personality qualities, and a battery of aptitude tests, indicating science-mathematical aptitude. The two major variables extracted and correlated by canonical analysis from these scores were measures of personality and measures of science-mathematical aptitude.

The results of the study reveal that correspondent Factor 1 components indicate that high computational and mathematical, with somewhat less general intelligence, correlates to a personality described as somewhat introverted and compulsive, but not possessing qualities of depression, schizophrenia, and asociality.

Factor 2 components relating scientific mathematical ability to personality patterns indicate that the person of high intelligence and high science ability tends to somewhat bizarre behavior combined with literary and musical interests, but is lacking in the qualities of accuracy with factual material and compulsiveness, indicating that the person with science aptitude and high intelligence may tend to "hang loose."

The results may have implications for counselors and vocational guidance personnel, if further research would bring to a more scientific level some of the commonplace stereotypes concerning scientists, and bring to an end some of those that do not apply.

◆ ◆ ◆

Involvement In And Apathy Towards Social Action Activities

DAVID W. JOHNSON and DANIEL C. NEALE, University of Minnesota

The attitudes and behavior of parents, the availability of models outside the family, and reference group norms have all been presented as determining conditions for prosocial involvement. In order to study the influence of these factors, a comparative study was made of a random sample of the undergraduate students at the University of Minnesota (N = 418; 82% response) and participants in a volunteer social action project (N = 84; 84% response). In a mailed questionnaire, respondents were asked to provide information about the social action activities and opinions of their fathers, mothers, friends (as a group), reference groups, and individuals (not members of their families) who had had a significant influence on their opinions and values.

The results indicate that the reference groups and "significant others" of the volunteers (compared with those of the sample) (1) have more favorable opinions about social action; (2) are more involved in social action activities; (3) have more influence upon the respondents' opinions; and (4) approve more of the respondents' participation in social action activities. In addition, the opinions of the volunteers' friends toward social action and towards the volunteers' participation in social action activities are more favorable and approving than those of the sample. Finally, compared with the sample, the fathers of the volunteers more highly approve of the volunteers' participation in social action activities, and the mothers have more favorable attitudes toward social action, are more involved in social action activities, and more highly approve of the volunteers' participation in social action activities.

◆ ◆ ◆

College Student Ratings of Instructors

NICHOLAS F. RAYDER, Michigan State University

The study was designed to answer two major questions : (1) do student sex, age, major, level of education, grade point average, or course grades previously received from the instructor being rated have any relationship to ratings of instructors? (2) are instructors who differ in sex, age, faculty rank, degree held, major area, or length of teaching experience rated differently by students?

Using a multiple regression analysis of data collected on 87 instructors from 4,285 usable student scales, the following conclusions were reached: (1) students remembered and accurately reported their grade point averages; (2) student ratings of instructors were not substantially related to students' sex, age, grade level, major area, grade point average, or grade (s) previously received from the instructor they rated; (3) generally, mean student ratings of instructors showed more relationship to instructor characteristics; (4) the correlations of the three student rating criteria with instructor age, faculty rank, highest degree held, and years of teaching experience, yielded twelve correlations varying from —.25 to +.15. Nine of these, however, were negative, suggesting that older, better trained, and more experienced instructors were not thought of as highly as younger and less experienced instructors; (5) mean student ratings of instructors according to major areas (departments) differed.

1:00 to 1:55 P.M. SESSION 18.6 DIVISION C

EVALUATION OF TEACHING STRATEGIES II

CHAIRMAN:

Lawrence T. Frase, Bell Telephone Laboratories

PAPERS:

Training Students to Tutor

GRANT V. HARRISON, University of California at Los Angeles and
RALPH J. MELARAGNO, System Development Corporation

There is increasing use of upper-grade students as tutors for younger children. A problem common to various methods of students serving as tutors is the training of tutors. The present study was conducted to develop training for fifth- and sixth-grade students who tutor first-grade learners.

Two approaches to this tutoring can be identified: structured tutoring, using pre-planned instruction with the tutor acting as a feedback mechanism and assisting learners with difficulties; and unstructured tutoring, in which the tutor chooses what to do with a learner and how. This study dealt with training for structured tutoring.

The pre-planned instruction was conveyed by booklets and tape recordings; taped messages called for the learner to make various responses to pictorial presentations in a booklet. The content was reading-readiness skills of particular necessity for Mexican-American children.

Seven basic skills for effective tutoring were identified. Training on the skills was developed empirically, using the trial-and-revision strategy: a tutor was trained by the experimenter; the tutor worked with a learner under the experimenter's observation; training was revised to correct deficiencies. Similar trial-and-revision cycles continued until tutors demonstrated proficiency with the seven skills; 10 tutors participated in these trials.

The revised procedure next was tried out with an elementary school teacher training four tutors, under the experimenter's direction. The four tutors demonstrated competence in most skills, and procedures were revised for implementation by naive tutors.

The procedure currently is being tried out in two elementary schools, where personnel are conducting training without assistance from the experimenter.

Results of these trials will be reported, as well as a description of the procedure and recommendations for training upper-grade elementary tutors.

Two Systems for Evaluating Videotaped Instruction

JACK GORDON, Oregon State System of Higher Education

The advent and ease of videotape replay has prompted another level of analysis of teacher-student interaction, a level that does not employ tedious dialogue transcription. Two separate evaluative systems have been developed for different studies. One involves a model of a structured, one-to-one "test" of teaching the concept "hexagon"; the other deals with determining changes in frequency of classes of behavior in samples from non-structured, "live", classroom interactions.

The structured "test", designed to differentiate between trained and non-trained EMR teachers, demanded that a model of "perfect" teaching be constructed as a standard from which to judge the videotaped performances. The detailed breakdown of the teaching task revealed that, 1) a concept such as "hexagon" is defined by a principle, 2) it is taught as a complex sequence of at least four psychomotor operations and a decision rule, all retrieved from memory, 3) the testing of pre-requisite skills is crucial, especially with EMR students. Intra- and inter-rater reliabilities have been in the .90's and reached the point where only one viewing of the tape is sufficient to complete an analysis.

The frequency change determination constitutes the evaluative procedure for a second study dealing with teacher self-evaluation using videotapes. Three 50-minute samples of classroom interaction have to be compared using only teacher intended change as direction. Characteristic interaction patterns and variability estimates are obtained for each teacher from the first sample which act as baseline data to judge whether the following samples do, in fact, differ. A simple set of categories was developed from which separate teacher flow patterns were classified. Three and five-minute sub-samples frequency counts of all possible combinations are made. Bayesian, information theory, and pattern analysis statistical methods are being tested for appropriateness.

Detailed manuals and scoring sheets for each system are to be appended.

◆ ◆ ◆

An Experimental Study of the Influence of Individual vs. Group Instruction on Spatial Abilities in Preschool Children

SISTER JOSEPHINA, Boston College

The increasing school enrollment of young children calls for a re-examination of preschool practices. Psychological findings on early childhood education stress the social and physical development as a function of learning. Concomitantly, Montessori devotees place important stress on academic accomplishment with the basic rationale of individualized instruction.

The proposed study investigated the possible differential impact of individual vs. group treatment of experimentally induced learning experiences involving principles of haptic-perception in the abilities of 3, 4, and 5 year old children.

The experimental group (N = 96) was equally divided by treatment: individual or group; the control group (N = 48) received group treatment. Each of the groups was further dichotomized by type of school enrollment: Montessori or non Montessori.

All Ss received five pre- and final tests given individually. During a 15-week period the Ss in experimental group received three 20-minute lessons per week. Those in individual treatment were taught in a one to one relation, while those in group treatment (6 Ss per group) received the identical 20-minute lesson given by some instructor three times per week.

The general statistical approach to the analysis of the data was a multiple analysis of covariance. Each of the five score variables represented change or growth in one aspect of haptic recognition, and served as the criterion variables of interest in a separate covariance analysis. Four types of independent variables, experimental condition, school, session attended, and sex were factors for which main effects and certain interactions were tested in the presence of the four independent variables—M.A., C.A., years in school, and personality factors. Fortran programs were used for intercorrelation, multiple correlation, and F-ratio.

In summary there resulted overarching concerns with answers to these problems.

1. What are the merits of group *vs.* a tutorial approach?

No difference resulted when data from Ss receiving individual treatment were compared with data from Ss receiving group treatment.

2. Do play experiences requiring active feedback from the learner in terms of verbalization, identification and matching of topographical forms, and graphic representation induce greater acceleration of learning than relatively free unstructured play situations?

All post-test scores differed significantly from pre-test scores. Ss receiving experimental treatment differed significantly from control Ss.

3. Can experimental verification be established for the assumption that tactual sensorial learnings from freely chosen media induce more perceptual learning than that found in free play or neutral activities?

Mean gain differential scores for Ss from Montessori classes deviated significantly from scores of the Ss in the non Montessori group. The Montessori curriculum with the emphasis on sensorial learning induced perceptual achievement.

1:00 to 1:55 P.M. **SESSION 18.7** **DIVISION B**

CURRICULUM DEVELOPMENT

CHAIRMAN:

Maurice J. Eash, Hunter College

PAPERS:

An Experimental Study in Curriculum Engineering

HARRIET TALMAGE, University of Illinois at Chicago Circle

The feasibility of describing the operations of the curriculum system through the use of a systems analysis model was tested experimentally. The problem was to determine whether concept attainment inservice training, Treatment A, significantly aided committees engaged in developing curricula. To test the problem experimentally, the effects of three treatments on three aspects of group performance were observed. The treatments included two types of inservice training (Treatments A and B), and no training (Treatment C). The data collected on group performance included: 1) the curriculum materials developed by committees; 2) the interaction behavior among committee members; and 3) the verbal content emphasized by committees.

Thirty *ad hoc* committees, composed of volunteer teachers, administrators, and graduate students, were randomly assigned to curriculum development tasks in one of three school subjects. Each committee consisted of three members. The three treatments and three school subjects, arranged in a 3×3 design, permitted observations in nine test situations. Each committee met separately for a three-hour time block. Following a fifty minute training for Treatment A and B committees, the task was assigned to each committee. The remainder of the time was divided into two sessions, one for discussion, the other for writing the curriculum.

A *Rating Scale for Curriculum Evaluation* was developed. The twelve-item *RSCE* quantitatively measured the task solution. Treatment A committees were rated significantly higher on the *RSCE* scores than Treatment B or C committees ($p < .01$). Committees receiving conventional training, Treatment B, did not perform significantly 'better' than committees receiving no training ($p > .05$). Differences among school subjects and the interaction between treatments and school subjects were not significant at the 5 per cent level of confidence. The treatments and school subjects had no significant effects on group interaction behavior. Only one category under verbal content emphasized (cognitive learning objectives) indicated significant differences between Treatment A and C ($p < .05$).

The use of the model in future studies would help in further sorting the variables of the curriculum system, controlling those variables not under examination, and testing selected input data for their value to curriculum committees.

High School Self-Evaluations and Curriculum Change

VYNCE A. HINES and WILLIAM M. ALEXANDER, University of Florida

This project sought to determine the scope and extent of curriculum change occurring in selected high schools as a result or concomitant of school self-evaluation conducted in accordance with regional accreditation policies based on the *Evaluative Criteria,* and of the impact of participation on selected attitudes and practices of the teachers involved. *Curriculum Change* meant any addition, subtraction, or modification of courses, activities, or services provided by a secondary school for its pupils.

A stratified, random sample of 60 Georgia and Florida high schools was used. Historical, interview, observation, and instrument data collection followed in pilot, experimental, and control schools to (1) validate and improve procedures, (2) find out if there were differences in amount of change in experimental and control schools, and (3) determine whether the amount of change was related to the pattern of preparation—no outside help, outside consultants, or faculty participation in a field labortaory course. Attitude changes among teachers in experimental and control schools were also studied as were certain teacher practices. Those teachers still in schools who were there at the time of the self-study responded to an attitude inventory on curriculum change.

Curriculum change occurs more often in the experimental schools during the year of the self-study and following year than in the control schools during the same period. Schools with no help make 25 percent more changes than their controls; those with outside consultants, 52 percent more changes; those with a field laboratory course, 77 percent more changes.

Curriculum changes are most likely to be course changes, faculty initiated, positively influencing attainment of school objectives, and to be judged as minor changes by the research staff. About 50 percent of the visiting committee recommendations are rejected. Changes are most frequent in English, then student activities, then instructional materials and library services, then guidance services.

Openness to change, teacher morale, and teacher activities did not change differentially either between experimental and control schools or among experimental schools by type of preparation.

To maximize the impact of school evaluations on curriculum change, evaluations should occur every five or six years and schools should arrange for field laboratory courses.

Behavioral Objectives and Test Item Production: Boon or Bane

GERALD E. GAGE, HAROLD D. FREDERICKS, BRADLEY HUITEMA, DONALD KOHL, and JAMES H. BEAIRD, Oregon State System of Higher Education

With the advent of programmed instruction has come an emphasis upon the necessity to state instructional objectives in behavioral terms. The utility of the behavioral objective for curriculum designers and measurement specialists is at the surface level inherently evident. The specificity of such objectives should simplify the task of the test item writer.

A difficulty arises, however, when one attempts to describe more complex objectives in specific behavioral language. More complex objectives involve more behavioral components, and result in greater disagreement about just which components actually describe the objective.

This study was concerned with one of the first problems involved in the establishment of a computer-based test development center—that of obtaining specific objectives and items to be used in the development of a large pool of instructional objectives and test items which would become the core of the total operation. Since relatively little is known concerning the optimum strategy for the training of item writers, it was felt that an experiment involving a limited number (60) of teachers (potential item writers) should precede any large scale item writing endeavor.

In order to obtain information concerning the most effective approach in item writing training, an experimental design utilizing three independent variables and six dependent variables was employed. Independent variables were:

(1) type of objective writing instruction (behavioral or not);

(2) type of item and objective writing unit (i.e. team, pairs, or alone); and

(3) source of objectives (i.e. whether the items produced are to evaluate objectives produced by self or others).

Dependent variables were:

(1) quantity of ojectives produced;

(2) quality of objectives produced (i.e. specific VS ambiguous);

(3) quantity of items produced; and

(4) quality of items (item-test correlation, content validity, and uniqueness).

The effects of the independent variables upon the above measures of quality and quantity of objectives and items were analyzed with six separate factorial design analyses or variance.

Several statistically significant differences were found. Of particular interest were the effects of objective writing training on the quality and quantity of objectives formulated and items produced.

The results suggest effective means of handling work teams engaged in the writing of items or specific purposes. More important issues, however, may be the overall effect of the item writing sessions on the teachers' classroom behavior, and changes in student achievement and attitudes resulting from changes in teacher behavior.

◆ ◆ ◆

Perceptions of Decision-making in Elementary School Classrooms

BERNICE J. WOLFSON, University of Chicago, and SHIRLYN NASH, Milwaukee Public Schools

Every day countless decisions—both major and minor—are made in every classroom. How do children and teachers perceive their roles in these decisions? Do children in the same class see pretty much the same opportunities for making their own decisions? Are these opportunities the ones their teacher believes she provides? Where disagreement exists, what kinds of decisions are involved?

The children in twelve elementary school classrooms (approximately 300 children in all) and their teachers responded individually to a fifty-item questionnaire. Some items concerned matters of classroom management (e.g., Who decides what desk or seat you can sit in? Who decides who cleans the blackboards?); others concerned the planning and scheduling of academic activities (e.g., Who decodes the plans or work for the day? Who decides when it's reading time?). For each item the subjects were asked to identify whether the child, the class, the teacher, or some one else was usually responsible for the decision. The same data were gathered twice, in October and again in May.

The responses of the children in each class were totaled and a mean frequency was obtained for each category. This frequency was compared with their teacher's frequency score in each category. Individual questionnaire items were analyzed for high or low agreement between the students and their teacher.

Analysis of the Fall data revealed a wide discrepancy between the perceptions of teachers and their students about who makes various decisions in the classroom. Teachers saw children making many more decisions than children saw themselves making. Teachers saw themselves making fewer decisions than the children saw them making. In the Spring administration of the questionnaire the discrepancy reappeared.

A closer look at this major finding was provided through an examination of the specific items on the questionnaire. Alternate explanations involving the requirements of institutional roles were explored. The relation of the findings to educational goals was discussed.

1:00 to 1:55 P.M. **SESSION 18.8** **DIVISION C**

CONSERVATION CONCEPTS II

CHAIRMAN:

B. Claude Mathis, Northwestern University

PAPERS:

A Model for Relevant and Irrelevant Transformations in Children's Concept of Weight

FRANK B. MURRAY and PAUL E. JOHNSON, University of Minnesota

A model, based on Newtonian mechanics, of the relevant and some irrelevant variables in the concept of weight was constructed to support a systematic investigation of the psychological model the child apparently uses in making identity conservation judgments. Three relevant transformations, i.e. those which changed the clay ball's weight, and six irrelevant transformations, i.e. those which changed some characteristic of the ball other than its weight, were studied. The relevant transformation were: (1) addition and subtraction of mass, (2) upward vertical change in position, and (3) downward vertical change in position; the six irrelevant transformations were: (1) change in shape of the ball, (2) division of the ball into three parts, (3) horizontal position change, (4) change in context in which the ball was placed near larger or smaller balls, (5) decrease in the ball's temperature, and (6) increase in the ball's temperature.

Eighteen items, two for each transformation, were randomly presented to a random sample of 120 second graders. In addition, children were randomly assigned to four groups in which were contrasted whether the items referred to weight as operationally defined by a beam balance or weight as unspecified, and whether items referred to a forward or reverse direction of the transformation. In the reversed transformation, the transformed state of the ball was presented first and then changed into the initial clay ball.

The analysis of the data indicated that all the "logically" irrelevant transformations, except horizontal position change, were relevant for the children. Only the first of the relevant transformations above was relevant for the children. Excellent reliability between the pairs of items for each transformation, and no overall balance—non-balance effect was found, as well as the theoretically significant result that transformations whose relevancy was misjudged were nevertheless reversible; that is, reversibility is not, as Piaget has often said, a sufficient support for conservation.

Conservation, Counting and Cardination as Factors in Mathematics Achievement Among First-Grade Students

GRAYSON H. WHEATLEY, Purdue University

This study was undertaken to investigate certain number concepts possessed by first-grade students. Specifically, the concepts of conservation (number and length), cardination, one-to-one correspondence, and counting ability were examined and their role in learning mathematics explored. A Piagetian-type number concept test (NCT) was developed and administered to a sample of entering first-grade students. As well as intelligence data, achievement test scores were available for the subjects.

Previous studies have shown that conservation of number is an extremely important principle in developing a concept of number and that children cannot understand addition of whole numbers without conservation. In light of these studies it was hypothesized that conservers would show greater achievement in studying mathematics than nonconservers.

The findings indicate that for first-grade students, achievement scores at end of year are highly related to scores on the number concept test administered upon entering school. Even when intelligence is partialled out, there is still a highly significant correlation between NCT scores and achievement scores. It was further revealed that there was a significant growth in number of concepts during the first-grade. Number conservation was the best single predictor of success in studying mathematics at the first-grade level of the factors considered. Counting was a poor predictor. It was found that fully one-half to two-thirds of the entering first-grade students did not understand number conservation and nine-tenths could not display knowledge of length conservation.

◆ ◆ ◆

An Analysis of the Effects of a Structured Teaching Approach Developed From the Work of Piaget, Bruner, and Others

HENRY P. COLE, State University of New York at Buffalo (Ronald Raven)

Inhelder and Piaget report that subjects younger than 13 or 14 years use 4 common irrelevant principles to explain why objects float or sink. A test requiring the selection of the correct principle from among the 4 irrelevant principles to explain why objects floated or sank in a series of demonstrations was developed. It was reasoned that a person who had truly attained the concept, which Piaget calls "specific gravity", would resist the strong "perceptual seduction" feature of each demonstration and maintain the use of the correct principle. The test is very similar in intent and purpose to the trick test used by Smedslund in his conservation of weight studies. The instrument was shown to be valid and reliable. Nearly all college junior and senior non-science majors studied, had not attained the concept in a stable form. The test was next administered to a group of 61 college students following their conventional instruction in the concept area. The criterion instrument showed this group had not attained the concept in stable form. It was reasoned that their poor performance could be due to the use of the 4 common irrelevant principles which the subjects had never learned to reject. A taped and filmed instructional sequence was developed based upon the reported work of Piaget, Bruner and others. The first part of the sequence was designed to force the rejection of the common irrelevant principles. The second part of the sequence was designed to develop the correct principle. A study was done on 38 non-science major, college juniors. Significant effects (.0001) of great strength of relationship were found to exist between the fully treated groups and the control group on the criterion instrument. However, no significant differences were found to exist between the groups receiving only the second part of the instructional sequences and a control group.

◆ ◆ ◆

A Comparison of Two Techniques for Using Visual-Tactual Devices to Teach Exponents and Non-Decimal Bases in Elementary School Mathematics

CECIL R. TRUEBLOOD, The Pennsylvania State University

An experiment was conducted to provide evidence on whether S's (age 9-11) would achieve and retain more by (1) manipulating visual-tactual aids (T-1), or (2) observing and telling the teacher how to manipulate such devices (T-2). Piaget's stages of intellectual development were used to hypothesize that T-1 would be superior to T-2.

Seven fourth-grade teachers volunteered to teach twenty-one lessons about exponential notation and non-decimal bases (EX-NDB) to their classes. These teachers followed structured lesson plans that specified:

(1) the visual-tactual aids used, (2) the basic inductive sequences followed,

(3) the amount of time spent on concept development and practice, and

(4) the content of the practice material.

The pupils in each class were randomly assigned to T-1 or T-2. An EX-NDB achievement test designed by the investigator was used to measure pupil achievement and retention. The EX-NDB test was given immediately following T-1 and T-2 to measure achievement and after a four-week period to measure retention.

An analysis of covariance was performed on the immediate post-test and retention test scores. Mental age was used as the covariate. The results of the covariance analysis indicated: (1) the pupils taught by T-2 scored higher on the immediate post-test than the pupils taught by T-1; the difference, however, was marginally significant ($p < .10$) and contrary to the hypothesis made from Piaget's theory; (2) pupils taught by T-2 did not retain significantly more than the pupils taught by T-1; nevertheless, both T-1 and T-2 resulted in a high degree of retention; (3) the EX-NDB achievement test scores had a high positive correlation with mental age. Finally, fourth grade pupils taught via T-1 and T-2 were able to learn and retain EX-NDB concepts presented by teachers who used the same basic inductive question sequences.

1:00 to 1:55 P.M. **SESSION 18.9** **DIVISION E**

CHARACTERISTICS OF COLLEGE POPULATIONS

CHAIRMAN:

Jack D. McFadden, Cornell University

PAPERS:

Pre-College Environment and Anticipated Role Orientations in College

ROBERT E. STANFIELD and HARRY SCHUMER, University of Massachusetts

The responses of 1600 entering students at a State university to a schedule of activities and behaviors available at college, produced through factor analysis eight basic orientations to the role of college student: vocational, instrumental collegiate, intellectual, consummatory collegiate, social development, ritualistic, academic, and fraternity-sorority. The factor scores of the students were analyzed with responses to a questionnaire eliciting information on family and secondary school experience. Chi-square analysis with the 25 percent highest scoring and 25 percent lowest scoring students on each of the eight factors, established certain significant associations between a student's background and his initial orientation toward being a college student. Among the relationships found were:

Vocationally oriented students tend to come from families of lower socioeconomic status, to have decided on a major field of study earlier than other students, and to be planning a professional life.

The instrumental collegiate orientation toward planning and organizing college activities tends to be particularly favored by Jewish females.

Intellectually oriented students tend to come from families of higher socioeconomic status, to be Jewish males, to plan on majoring in the humanities, and to express an interest in a life that is academic or related to creative arts.

The consummatory collegiate orientation toward enjoying the benefits of the collegiate culture is favored by students who see their future life in a technical field.

Females majoring in social sciences tend to be high on the intellectual orientation and low on the academic, but females majoring in physical sciences tend to be low on the intellectual and high on the academic.

Expressing a desire to pursue graduate study after college tends to be associated with high scores on vocational, instrumental collegiate, intellectual, and academic, but with low scores on consummatory collegiate and ritualistic.

Factors Which led College Seniors to Choose College Teaching as a Career

JACK E. ROSSMANN, University of California, Berkeley, and JOSEPH C. BENTLEY, University of Utah

In an attempt to understand better the factors which motivate college Seniors to seek careers in college teaching, a questionnaire and the *Strong Vocational Interest Blank* were mailed (in February of 1966) to the 97 men who were awarded Danforth Graduate Fellowships the preceding fall. They had applied for the Fellowships and, in effect, decided to head toward careers in college teaching in the Fall of their senior year in college. Completed questionnaires and SVIB's were obtained from 94 of the 97 Fellows.

The major findings were:

1. The Fellows reported having chosen teaching for reasons of intellectual challenge, desire to work with college-age students, and the pleasant atmosphere in which to work. The most frequently selected response was, "Felt I could make the greatest contribution to society in this area."

2. 15% of the sample had decided to become college teachers during their high school years, and one-half the Fellows had reached that decision by the beginning of their junior year in college.

3. When asked the satisfactions they anticipated from college teaching, 90% of the Fellows mentioned a satisfaction related to the collegiate environment or personal associations. Other anticipated satisfactions included opportunities for service to others, self-development as scholars and teachers, and satisfactions from the teaching process itself.

4. Primary areas of anticipated dissatisfaction included pressures to publish; the burden of administrative duties and relationships with administrators; and problems resulting from competition for status and recognition.

5. The Fellows anticipated that they would spend an average of 55% of their professional time in teaching activities in their first position. Twenty years later they expected to reduce this to approximately an anticipated 40% on the average, and anticipated spending more time in research and writing (30%), committee work (7%), off-campus consultation (8%), and administrative duties (8%).

6. As measured by the Strong Vocational Interest Blank, the interests of the Fellows most closely resemble those of physicians, psychiatrists, psychologists, and biologists; librarians and musicians; and lawyers and writers.

7. 84% of those Fellows who had done their undergraduate work at a university hoped to return to the university setting to teach, while only

53% of those who did their undergraduate work at a liberal arts college hoped to teach at a liberal arts college.

◆ ◆ ◆

Who Goes Where to College? A Cross-Validation

JAMES M. RICHARDS, JR., University of Calfornia, Los Angeles and
LARRY A. BRASKAMP, University of Nebraska

In order to plan intelligently for the future of higher education, it is important to have reliable information about the allocation of students among colleges. Recently Astin published a large study of the relationship between college environments and student body characteristics. However, Astin's data are several years old, higher education is changing rapidly, and more comprehensive descriptions of college student bodies are available. The purpose of the present study is to relate more recent and comprehensive information about entering freshman classes to Astin's measures of college environments. Therefore, it is essentially a cross-validation of his environmental measures.

Astin provides eight environmental scores: size, selectivity, and six "personal orientations"—Realistic, Intellectual, Social, Conventional, Enterprising, and Artistic—based on the proportion of students in each of six classes of major field. Information concerning the characteristics of entering freshmen were obtained from the assessment administered by the American College Testing Program. This assessment provides a comprehensive set of information about entering freshmen: test scores, high school grades, special interest, campus needs, and non-classroom accomplishments. Information is provided both for enrolled students and non-enrolled applicants; that is, students who had ACT scores sent to that college but did not enroll.

For 179 four-year colleges, correlations were computed between this description of entering freshmen and Astin's environmental measures. Freshman characteristics covary in interesting ways with the environmental characteristics, and the pattern is both meaningful and generally consistent with the interpretation given the environmental measures. These conclusions are true only in a broad sense, however, for most of the correlations are moderate to low. Therefore, the environmental measures are not a completely satisfactory substitute for a detailed description of the student body. Results for non-enrolled applicants suggest that student body characteristics are determined more by a college's constituency than by its selection procedures.

Why They Came to the East-West Center: A Comparison of Asian-Pacific and American Grantees

GERALD M. MEREDITH, University of Hawaii

The East-West Center was founded to promote better relations and understanding between the United States and the nations of Asia and the Pacific through cooperative study, training, and research. As a joint venture of the University of Hawaii and the Federal Government, the guiding principle of the Center is interchange—the creation of a climate to enhance dialogue between college students of diverse cultural and linguistic backgrounds. Stdents are selected not solely on academic and degree-getting skills, but on ability to profit from an international and innovative educational experience.

The present research was concerned with the identification of major components in reasons given by Asian-Pacific and American grantees for the decision to attend the East-West Center. Specifically, it was hoped to (a) identify major factors in reasons given for attendance; (b) clarify the self-reported motivational structure behind educational-vocational decisions; and (c) examine the relationship of these factors to areas of satisfaction and dissatisfaction within the program.

Dole's *Reasons for Going to College* inventory was modified and administered to 216 Asian-Pacific and 106 American grantees enrolled at the University of Hawaii. The 93-items yielded pattern differences between the two student groups in the areas of Reasons, Values, Interests, Influences, and Needs determining educational-vocational choice. In addition, a 127-item questionnaire was used to glean information about grantee perception of ten facets of the program (e.g., nature of interchange experience, personal-social adjustment, academic programs and services, community relationships, etc.).

The findings of the study suggest two broad styles of "coming to terms" with a multi-cultural academic environment, and these were discussed with the reference to the motivational-vocational models of Tyler (stress on individuality) and Super (work values). Implications for educational decision making and for future research were considered.

2:00 to 2:55 P.M. SESSION 19.3 DIVISION C

VERBAL BEHAVIOR VI

CHAIRMAN:

Orpha K. Duell, Wichita State University

PAPERS:

Experimental Conditions Effecting the Performance of 5 Year Olds on a Letter Discrimination Task

VERNON C. HALL, EDWARD CALDWELL, GWENN SIMPSON, Syracuse University (Eric F. Gardner)

The purpose of this study was to determine the influence of motivation, warmup condition, letter size, instructions, and feedback on the performance of 5 year old children on a letter discrimination task. One hundred children were randomly assigned to 5 treatment groups and administered a letter discrimination task similar to the one typically used to determine readiness for letter training. One group was given the identical instruction given by an earlier study (Davidson, 1935), while other groups were given different warmup and motivation conditions (i.e. presenting candy for correct performance). In addition, within each condition, half of the children were given letters from a primary typewriter, and half were presented pica letters. The group identical to the earlier study performed almost exactly the same as reported originally. Analysis of variance indicated a significant improvement by the other conditions. Next, half of the Ss were given feedback as to how well they did (no training or cues were mentioned). An analysis of covariance indicated that feedback again significantly improved the performance of the children. Evidence is offered to show that the important variable is the child's definition of same and different rather than perceptual capacity of the children. When schools make decisions based on readiness tests not taking these considerations into account, much time and effort will be wasted.

◆ ◆ ◆

Effects of Incidental Cues and Encoding Strategies on the Learning of Vocabulary Equivalents

JOSEPH R. JENKINS, University of Delaware

The present study sought to determine the characteristics of pictorial representation that produce faster learning with picture stimuli than with word stimuli.

An incidental cue hypothesis postulated a difference in perceptual richness for words and pictures. Pictures were thought to be perceptually richer stimuli due to a salience of incidental cues, that is, cues that can be altered without changing the tendency of subjects to attach the appropriate label to the stimulus.

If incidental cues have been responsible for the observed differences in learning, then a condition in which incidental cues were washed-out from pictures (PWa)—different pictures were used to represent the same concept —should be comparable to a condition in which normal words (WN) were employed as stimuli. Secondly, a condition wherein incidental cues were added to words (WE) should be superior to the WN condition.

A second hypothesis postulated a difference in encoding processes for the two types of stimuli. There may be a weaker tendency to encode a stimulus verbally when it is a picture rather than a word, which in turn may reduce associative interference by decreasing responses elicited by the verbal label. This hypothesis predicted that requiring subjects to label the stimuli overtly should produce a decrement in performance, particularly with picture stimuli. Here a 2 × 2 design with labeling and stimulus mode composing the dimensions was employed.

Although results revealed the previously observed difference between picture and word stimuli (PL.01), group PWA unexpectedly performed as well as a *picture normal* condition (only one picture to represent each concept) and significantly superior to condition WN (PL.001). Also, enriching words by adding incidental cues did not facilitate learning relative to WN. Furthermore, requiring *S*s to label the stimuli overtly did not effect learning, nor did this interact with stimulus mode.

◆ ◆ ◆

Verbal Versus Graphic Presentation of Learning Material

PATRICIA J. PARSONS, University of Massachusetts, and LAWRENCE T. FRASE, Bell Telephone Laboratories, Murray Hill, New Jersey.

Limited research has been done on the effect of verbal and graphic methods of instruction—most of it indicates that *S*s profit equally from both types of material. There has been no systematic attempt to insure that there is a 1:1 correspondence between the two types of material and usually the verbal material offers some additional cues for learning. The present study tried to achieve such a correspondence, and to correct other ambiguities of the earlier research.

64 undergraduate psychology *S*s were divided into four groups of 16 each which saw the verbal or graphic materials alone, or different sequences of the two, i.e., verbal-graphic or graphic-verbal. The subject matter was electrical circuitry. The verbal material was programmed and then translated directly into graphic form, eliminating only connectives and necessary queries.

The experiment was administered in a large auditorium. Immediately after studying the materials, *Ss* took a test comprised of recognition, recall, verbal, and graphic subtests.

Analyses of variance indicated that recognition scores were higher than recall, but there were no differences between the experimental groups. The graphic groups tended to be slightly better on all tests except the recognition items. A non-verbal culture fair test (Cattell & Cattell) had moderate but significant correlations with both verbal and graphic performance ($r=.33$, $p<.01$). Total time spent on the material also correlated significantly ($r=.35$, $p>.01$) with overall performance, as well as with individual subjects. Subjects stated that when they encountered graphic materials, they first translated them into words, then solved them. There was no relation between time spent and SAT verbal ability, hence such translation probably did not differentially benefit high verbal ability *Ss*.

Students evidently learn equally well from coordinate verbal-graphic methods of presentation.

◆ ◆ ◆

Listening Comprehension as a Function of Word Rate

EMERSON FOULKE, University of Louisville

The purpose of the study was to clarify the relationship between word rate and listening comprehension. Word rate was varied by the Tempo Regulator, a speech compression device that varies word rate by systematically discarding or repeating samples of recorded speech. A listening selection was presented at 12 word rates, starting with 125 wpm, and increasing in steps of 25 wpm to 400 wpm. Listening comprehension, as measured by a multiple choice test, remained fairly constant through 250 wpm, and began to decline at an accelerating rate thereafter. The lack of a linear relationship between word rate and listening comprehension is discussed in terms of its implications for the cognitive processes that support the perception of spoken language.

2:00 to 2:55 P.M. **SESSION 19.4** **DIVISION D**

STUDIES OF PROSPECTIVE TEACHERS

CHAIRMAN:

Paul J. Blommers, State University of Iowa

PAPERS:

Personality Patterns and Heterogeneity Among Teaching Interest Areas: A Multivariate Analysis Approach

ERNEST W. BODENSTAB, Valparaiso University

The study sought to demonstrate that an association exists between the personality of the student teacher and the specific teaching interest area selected by the student. Such a demonstration is appropriate since no resolution has been made concerning psychological homogeneity or heterogeneity for the teaching profession.

The data consisted of 571 Minnesota Multiphasic Personality Inventory profiles of student teachers. Multivariate analysis of variance determined the existence of psychometric differences among the grade level groups of elementary, special, and secondary. A classification procedure compared the differences with the subject's personality pattern and assigned each subject to one of the three groups. The multivariate analysis of variance analyzed the 571 subjects for psychometric differences when subgrouped in eleven teaching interest areas: (1) elementary, (2) Health, Physical Education, and Recreation, (3) art, (4) music, (5) speech, (6) English, (7) social studies, (8) mathematics, (9) languages, (10) science, and (11) business education. The classification procedure assigned each subject to one of the subgroups. Each analysis recorded the number of subjects classified in agreement with the subject's choice. After chance probability assignment was subtracted, a comparison of the analyses' results sought a greater percentage for subgroup classification than for group classification. The greater than chance assignment to groups (three grade levels) was 17% and the greater than chance assignment to subgroups (eleven teaching interest areas) was 33%.

The successful demonstration strengthens the following positions and provides direction for future research. First, the teaching profession is psychologically heterogeneous with respect to teaching interest areas. Research seeking communalities for the teaching profession must account for the heterogeneity. Second, the research design employed multivariate analysis of variance which evaluated twelve variables of a psychometric instrument. The study utilized an original method for comparing the results of the analyses when applied at a group and subgroup level.

Observation and Evaluation of the Classroom Behavior of Student Teachers

BOB BURTON BROWN and JAMES T. McCLAVE, University of Florida

The purpose was to identify variables and combinations of variables which contribute significantly to variance in observations and evaluations of the classroom behavior of student teachers.

A total of 569 observer-judges made 2859 observations and 953 evaluations of 407 student teachers from six teacher education institutions in California, Florida, Illinois, New York, and Wisconsin. In addition, scores on the *Personal Beliefs Inventory,* the *Teacher Practices Inventory,* and the *Dogmatism Scale* were obtained for each student teacher and observer-judge in an effort to measure the influence of the personal and education philosophy, and the structure of belief systems on the observations and evaluations.

The data were submitted to multiple regression analysis to discover what measured variables contributed to the scores given on the *Teacher Practices Observation Record* (TPOR) and the *Teacher Evaluation Scale* (TES), yielding many interesting findings with respect to variables (and factors within variables) which are predictive of observation and evaluation scores.

If meaningful observations and evaluations of the classroom behavior of teachers are to be made on a large scale by untrained personnel, the findings of this study indicate that information regarding the uncontrolled beliefs and biases of the observer-judges, as well as those of the teachers, is essential.

◆ ◆ ◆

Use of Discriminant Analysis in a Study of the Patterns of Characteristics of Mathematics Teachers

JUNE R. CHAPIN, College of Notre Dame

The statistical procedure of discriminant analysis was used in a study of 177 mathematics student teachers of a state college. Discriminant analysis, a procedure from multivariate analysis, was used to obtain weights for 40 characteristics such as married, grade point average, professed avocational interests, supervisor's ratings during student teaching, etc. These weights were applied to compute a total score for the individual's characteristics and this score was used with the criterion of teacher effectiveness, the principal's rating of their present-day teaching, which had been obtained from a follow-up procedure on 95 of the student teachers. After extensive sorting through 40 variables, it was found that only a few contributed significantly to the score but discriminant analysis had enough power to distinguish the patterns of characteristics of four groups (successful, unsuccessful, employed by school district but not teaching mathematics, and those who had left teaching) at the .01 level of significance.

Some potential areas of use of discriminant analysis in educational research are given.

◆ ◆ ◆

Changes in Pupil-Teacher Attitudes: Yeasayers vs. Naysayers

DONALD L. LANTZ and WM. K. BOTT, University of South Florida

This study was designed to investigate changes that occurred in students' attitudes of pupil-teacher relationships as measured by the MTAI among high and low scorers on the F scale.

The subjects were 192 students that were enrolled in their first education course. Subjects were in one of eight sections taught by one of three instructors. The organizational structure of the course consisted of one lecture and two discussions per week. The *Ss* were classified by two factors, initial MTAI score and F score. Each of the factors consisted of two levels. The two levels for each of the factors corresponded to the upper and lower halves of the distribution on the two instruments.

The MTAI and F scale were administered to the subjects on the first day of class and the MTAI was administered again on the last day of class.

The data were analyzed by two-way analysis of covariance. An .01 level of significance test indicates significant differences between the two levels of the F scale. No significance was found between levels of initial MTAI score or interaction of initial MTAI score and F score.

The data clearly indicate that when initial attitudes and intelligence are statistically controlled, beginning students in professional education who have low F scores are more amenable to education in the sense they make more positive changes in attitudes as measured by MTAI than those who have high F scores. Students who may be characterized as independent and critical (low F scorers) actually show greater MTAI gains than students who have a tendency to agree with any assertion or who hold antidemocratic attitudes (high F scorers) in classrooms characterized by a high degree of classroom discussion.

2:00 to 2:55 P.M. **SESSION 19.5** **DIVISION D**

EMPIRICAL TEST RESEARCH

CHAIRMAN:

Norman Frederiksen, Educational Testing Service

PAPERS:

An Empirical Investigation of Item Sampling

TED HUSEK and KEN SIROTNIK, University of California, Los Angeles

This paper reports two attempts at empirical verification of the concept of item sampling. The first employed a population of 200 students and their responses to a 140 item final examination on the course objectives. The second attempt used a population of 350 students and their responses to a 30 item mathematics achievement test. Each population of students and items was randomly divided into ten, nonoverlapping item samples. Traditional examinee samples were also formed by combining every two item samples and considering the responses of each student on all the items. Estimates of the population mean and variance were computed for each of the item and examinee samples. Finally, comparisons were made between the "average" item sample's estimate of these parameters and those of each examinee sample.

Results support item sampling as an empirically useful technique for estimating the mean and variance of a population of scores. The item sampling procedures were more efficient than the examinee sampling methods without exception for the second population. However, the data for the first population were not so dramatically favorable. Furthermore, one of the item samples produced a negative variance estimate. These two findings led to a careful re-examination of (1) the purposes for which the test was used in each of the populations and (2) the item sampling formula for the estimation of the variance. It was subsequently shown that (1) a negative variance estimate would be obtained when coefficient alpha of the item sample was negative and (2) a test designed to measure the attainment of objectives (as in the first population) would be more likely to yield such an item sample than a test designed to maximize variability among subjects (as in the second population).

On the basis of these results, implications for the practical application of item sampling methodology are discussed.

◆ ◆ ◆

Differential Validity of Factors in the Eisner Art Attitude Inventory

JOAN M. LOYKOVICH and P. R. MERRIFIELD, Kent State University

Eisner (1965, 1967) has reported results of the use of an inventory he constructed to obtain information about students' attitudes toward various aspects of art. He has reported the 60 items in terms of four separate categories: "voluntary activity in art"; "satisfaction in art", "self-estimate in art"; and "attitude toward art and artists." No report of any attempt to determine the statistically independent dimensions of the inventory has been made.

A revised form of the inventory was developed, with Eisner's permission. In the revised form, the items each written as a declarative statement, were rearranged so that one representative of each of the four different categories appeared in each block of four items.

In May, 1967, responses to the revised 60-item inventory were obtained from eighty-five tenth-grade boys and girls enrolled in a high school in northeastern Ohio. This group appears to be a sample of the same population as those Eisner has reported testing.

On the basis of a preliminary factor analysis of the 60 items and a profile of the eigenvalues, it appeared that seven factors would be sufficient to explain the interitem correlations. For the main factor analysis, twelve items were deleted on the grounds that they would contribute very little to the differential validity of the seven factors expected. In the main analysis the highest correlation of an item with any other item was used as the initial estimate of its communality.

Five clearly interpretable orthogonal factors were found. The first factor, "voluntary activity in art," is characterized by items which were in two of Eisner's categories, "voluntary activity in art" and "self-estimate in art." The second factor, "concept of self as artist," is characterized by items which are all in the Eisner category, "self-estimate in art," and are items from that category that do not contribute heavily to the first factor. Items on Eisner's "attitude toward art and artists" contribute to two factors, "acceptance of art as a social value" and "attitude toward artists." The fifth factor, "appreciation of art," is represented by items on Eisner's category, "satisfaction in art." With few exceptions the items are univocal. The sixth and seventh factors are not clearly interpretable.

◆ ◆ ◆

The Effect Upon the Factor Structure of Semantic Meaning of Various Sources of Variance

KEITH A. McNEIL, Southern Illinois University

Several sources of variance have been analyzed in factor-analytic work with the semantic differential. Only one source (that of scale variance) seems to be appropriate in determining the underlying dimensions of meaning. Typical semantic factor structures have been determined by also considering concept variance and the variance due to the interaction of scales and concepts. Subject variance may be more crucial than heretofore recognized and so one aspect of subject variance was also investigated.

521 Austin, Texas, sixth graders rated 12 concepts on 20 scales. The subjects were members of the following four subcultural groups: Middle-Class Anglo, Lower-Class Anglo, Lower-Class Negro and Lower-Class Latin-American. Two different approaches were used to generate the correlation matrix. Approach A averaged the scale ratings across the concepts, and then correlated over subjects. Approach B treated each S's responses to a single concept as a separate source of data. All correlation matrices were factored by the principal axes method with Varimax rotation of all factors having an eigenvalue above 1.0.

The results indicated that Approach B yielded a factor structure more consonant with previous research than did Approach A. The separate factor analyses of the various subcultural groups indicated that the factor patterns changed from one subject group to another. The changes were more drastic when Approach A was utilized than when Approach B was employed.

Concept variance and subject variance seem to be important sources of variance in factor analytic work with the semantic differential. Researchers are cautioned to check carefully the procedures followed and the subject population utilized before scales are borrowed from previous research.

◆ ◆ ◆

An Analysis of Subtest Performance by Gifted Students on the Stanford-Binet Intelligence Scale (1960 Form L-M)

ROBERT L. SHEVERBUSH, JR., School District 11, Colorado Springs, Colorado

The major purpose of this study was to analyze the performance of gifted children on the *Stanford-Binet Intelligence Scale (1960 Form L-M)* in terms of individual subtests and functional groupings of subtests. The writer investigated the problem by studying the responses made on the Stanford-Binet by 192 gifted children who had a mean chronological age of nine years eight months when tested and who were members of Highly Academically Talented Classes in an urban school district in Colorado. The responses of this group were compared with the responses made by a group of 192 children of nearly the same chronological age who scored in the average intelligence range. Sub-groups of the gifted were compared on the basis of teacher ratings as to degree of success in the gifted classes, on sex and on socioeconomic status.

The system of classification into functional categories that was used in this study was that of Carol L. Lutey (1966). This classification consists of: A. Manipulation Tests; B. Vocabulary and Language Tests 1. Vocabulary Tests 2. Language Comprehension Tests 3. Language Usage Tests; C. Memory Tests; D. Discrimination Tests 1. Perceptual Discrimination Tests 2. Verbal Comparisons Tests; and E. Reasoning and Problem

Solving Tests 1. Verbal Reasoning Tests 2. Nonverbal Reasoning Tests 3. Numerical Reasoning Tests

When Kendall's coefficient of rank correlation was applied to the differences in rank order of percents passing the subtests in the various functional categories the results were positively significant ($p = < .05$). The successful group had the highest rank on subtests of the reasoning categories; whereas the unsuccessful had the highest rank in the verbal categories. The gifted male group had a higher rank difference on subtests classified into the reasoning categories and the female group had a higher rank difference on subtests classified as verbal. The high socioeconomic status group had the highest rank on subtests of the reasoning categories; whereas the middle and low socioeconomic status groups had higher ranks in the verbal area.

The use of this classification system is helpful in organizing responses on subtests of the *Stanford-Binet* into meaningful patterns. It appears from the responses on individual subtests that there is consistency within each functional category.

2:00 to 2:55 P.M. SESSION 19.6 DIVISION E

VOCATIONAL INTERESTS AND CAREER DEVELOPMENT

CHAIRMAN:

Douglas D. Sjogren, Colorado State University

PAPERS:

Effects of Participation in a Community Resources Program on Selected Attitudes of Adolescents

RICHARD M. CLARK and REUBEN R. RUSCH, State University of New York at Albany

A three-year study was conducted to find the effects of participation in a Community Resources Program on the attitudes of adolescents. The study examined the relationships between the student's seminar experiences where he worked directly with a professional person, and his attitudes toward the professional field.

Students in the program and a selected group of control students were asked to sort sixty statements through a Q-sort procedure designed to reveal attitudes toward artists, writers, scientists, and social scientists. Also, the statements were sorted by two professional groups (artists and scientists), and their responses were compared with the responses of high school students.

A brief summary of the results of the study indicates that:

1. *high school students in general have a favorable view towards the professions;*
2. *high school students in general have a more favorable view toward science and social science than they do toward art and writing;*
3. *students with seminar experiences with artists viewed art more favorably than control students or students with other kinds of seminar experience;*
4. *students place more stress than professional scientists on the need for very high cognitive ability to succeed in the field of science.*

◆ ◆ ◆

Physics, Femininity, and Creativity

HERBERT J. WALBERG, Harvard University

Girls get higher grades in school even in subjects in which boys score higher on standardized tests. Yet, after graduation, men produce more creative works in science and the arts. This exploratory research focuses on a crucial time for career planning—the last two years in high school—and attempts to account for the discontinuity by examining the differences in abilities, motives, and outlooks of boys and girls studying physics, a presumably "masculine" and therefore conflict laden field for women.

Using a randomized data collection, 58 measures (IQ, physics achievement, science understanding, science activities, values, affiliation, change dogmatism, semantic differential ratings, and perceptions of classroom climate) were taken on 2100 students in 76 classrooms scattered throughout the nation. T-tests revealed significant differences ($p<.05$) between boys and girls on 36 measures, and omega squared statistics brought out the relative importance of sex differences in accounting for variance.

Four patterns of measures discriminate between boys and girls, and girls scored higher on all four: verbal aptitude (but not quantitative and spatial abilities), social values and affiliative needs, cautiousness, and aesthetic valuations. Viewed in the context of previous research, these factors may lead to greater academic success in high school, but appear to penalize women for later career eminence especially in science.

◆ ◆ ◆

Quantification of Subjective Data Via Content Analysis of Essays

VIVIAN S. SHERMAN, American Institutes for Research in the Behavioral Sciences

In recognition of the need for instruments and approaches designed to tap subjective as well as objective data and as part of a larger study, content analysis of essays was undertaken to assess the extent to which junior high school students appear to be involved in career exploration, and how they perceive themselves and their development relative to vocational interests. Specifically, a system of scoring verbalized self-perceptions and specificity-vagueness of career goals and plans was empirically derived.

Essays on career interests, gathered from junior and senior high school students, were used in preliminary establishment of 25 categories subsumed under three broad classifications: Personal Resources and Characteristics, Past Influences on Personal and Career Development, and Resources and Opportunities for Future Development. This initial scoring system then was employed as one pretest measure in a study designed to test the effect of curriculum experiences on students' attitudes. Application of the scoring system to essays of this larger sample of approximately 200 students resulted in some refinement of both category headings and illustrative examples. Procedures for establishing and reporting inter-rater reliability currently are being developed. Posttest application of the scoring system has been is set.

A method for quantifying subjective data involving the crucial variable of self in relation to career development will be presented and evaluated for practicability. A system for content analysis of essays will be described, and evidence provided regarding its use both as a research tool and for assessing individual differences. Problems encountered in its usage will be delineated, i.e., scoring time, reliability for two applications, discreteness and range of categories, areas needing further refinement, etc.

The scoring system described appears to hold promise both for ongoing school district research and providing guidelines for individual and group counseling, cues for curriculum planning, and normative data in vocational development.

◆ ◆ ◆

A Cubistic Model of Vocational Interests

AYRES D'COSTA and DAVID W. WINEFORDNER, State Department of Education, Columbus Ohio

The purpose of this paper is to explain the cubistic model of vocational interests, and present the results of studies conducted to test its efficiency as a model.

The data-people-things structure presented by the Dictionary of Occupational Titles (1965) defines a three-dimensional space which contains every job in the world of work. When the 114 worker trait groups defined by the D.O.T. were plotted in this cube, several clusters were observed. These clusters were found to be also homogeneous in terms of other job characteristics such as interests, temperaments, specific vocational preparation, and general educational development.

It is believed that a vocational interest survey constructed so that its scales correspond to these job clusters has the following advantages over existing instruments: 1) Its 'domain' or 'scope' is comprehensive of the world of 'vocations' and therefore of the world of vocational interests; 2) its scales, being defined in terms of worker trait groups, are more readily interpretable to counselees; 3) it lends itself well to vocational orientation and exploration using existing resources like the D.O.T. and its related literature on occupations and educational training programs.

The Ohio Vocational Interest Survey (1967) was built on these premises. Empirical findings to date include item-analyses, homogeneity, reliability, and validity of the 24 scales, as well as a multivariate discriminant analysis to compare six vocational education areas as defined by the U. S. Office of Education (1967). Computerized methods have been used in the analysis of the instrument and in the reporting of the results to counselees.

The study indicates that the cubistic model has considerable merit in terms of providing meaningful and effective scales for interest measurement. The authors caution that their approach departs from Kuder's approach of defining independent and factorially 'pure' scales. Their 24 scales have vocational significance like Strong's, but have the added advantage of being comprehensive in coverage.

3:00 to 3:55 P.M. **SESSION 20.1** **DIVISION B**

SCHOOL INTEGRATION

CHAIRMAN:

Ronald Hyman, Rutgers University

PAPERS:

The Effect of Race on Peer Evaluation and Preference in Primary Grade Children

SANDRA R. COHEN, and MARIANNE AMAREL, Educational Testing Service, Princeton, N. J.

The purpose of this study was to explore the attitudes of students in the primary grades wih regard to the learning ability, school behavior, and desirability as classmates of white and Negro children.

Subjects were 120 first and second graders at an all white, an all Negro, and an integrated school. In individual interviews, S's were shown sketches of classroom events taking place in predominantly white and predominantly Negro classes, and were asked to make judgments about the academic achievement, niceness, school behavior, friendliness, and general desirability of the children in the pictured classes.

The sketches were shown in pairs to the subjects. The two pictures in any pair were similar in all respects except their racial composition: one sketch in each pair showed a predominantly white class, while the other showed a predominantly Negro class.

Subjects were asked nine questions, including "which class do you think reads the best," "which class do you think has the nicest children," "which class do you think fights the most," "which class would you like to be in?" Subjects responded by pointing to either the predominantly white or predominantly Negro class.

Questions and picture pairs were rotated randomly and independently of each other. Subjects were retested following a four-week interval. Test-retest reliability was .63.

Data were analyzed with a four-way analysis of variance (race × sex × grade × integration). There was a highly significant effect for race ($p. < .0001$). White subjects showed a striking preference for predominantly white classes. Negro subjects showed no significant preference for either white or Negro classes.

There was no significant effect for integration. No differences were found between scores for Negro subjects in the integrated school and in the all-Negro school, nor between scores for white subjects in the integrated school and in the all white school.

The Effect of the Curriculum on Racial Cleavage in Integrated Fourth Grade Classrooms

SUSAN BAHLKE and P. JOHN GEORGEOFF, Purdue University

This study seeks to investigate what effect, if any, the school curriculum has upon racial cleavage in integrated fourth grade classrooms.

Sixteen integrated fourth grade classrooms were selected for the study from a large, metropolitan Indiana school city. Each of the classes selected had an enrollment in which at least 25% of the children were of one race. Nine of these classes served as the experimental group and seven served as the control group.

A sociometric device was administered to each of the classes as a pre-test. Each child was asked to make three choices from the children in his room with whom he would like to study, sit, and play; that is, in three-areas—academic, social, and athletic. Each child thus made nine choices altogether. Each of the experimental classes was taught a unit on American Negro history and culture for a period of six weeks. The unit, consisting of both factual and conceptual context, was taught five days each week for thirty minutes. At the end of this period, the sociometric device was again administered to both groups—experimental and control.

McNemar's test was then used to determine if any significant differences developed after study of the unit. The findings indicated that a significant shift had taken place from same to different in the case of the play and study, but not in the case of sit, the social aspect. In the first instance, the difference was at the .05 level of confidence and the second, at the .01 level.

It would appear, therefore, that a study of the cultural and historic contributions of the Negro to American life can reduce cleavage in integrated classes, in terms of at least two inter-personal relatinships—the academic and the athletic. The social arena of these relationships, however, being circumscribed by strong emotional family and community taboos, pressures, and restrictions, is much more resistant to change, and apparently cannot be effected over a relatively short period of time by brief, though concentrated, study of the topic.

◆ ◆ ◆

The Effect of the Curriculum upon the Self-Concept of Children in Racially Integrated Fourth Grade Classrooms

P. JOHN GEORGEOFF, Purdue University

This study seeks to ascertain what effect, if any, the school curriculum has upon the self-concept of children in racially integrated fourth grade classrooms.

Twenty-six integrated fourth grade classrooms were selected for the study from a large, metropolitan Indiana school city. Nine of these classes were composed of children who attended a school in their own neighborhood. Nine other classes included children who were transported to their school from other areas in the city because of over-crowded conditions in their own district. The remaining seven classes served as the controls.

The Piers-Harris Self-Concept Scale was administered to each of the classes as a pre-test. Each of the experimental classes was then taught a unit on American Negro history and culture for a period of six weeks. The unit, consisting of both factual and conceptual content, was taught five days each week for thirty minutes. At the end of this period, the Piers-Harris Scale was again administered to all three groups—experimental and control.

Analysis of co-variance was then used to determine if any significant differences had developed in the self-concept of the children in the experimental groups as a result of studying the unit. The findings indicated no significant differences to have taken place when the groups were taken as a whole, or when each group was divided into sub-groups accordng to race.

It would appear, therefore, that a limited study of the cultural and historic achievements of the American Negro has no measurable effect upon the self-concept of Negro or white children of this grade level. The factors contributing to self-concept formation, such as home, community, contacts, socio-economic standing, etc., appear to be too strong to be overcome in the relatively short period of time, by brief, though concentrated study of a unit of school work.

◆ ◆ ◆

School Racial Composition and Teacher Influence

JOSEPH H. McMILLAN, Grand Rapids, Michigan Public Schools; EDSEL L. ERICKSON, Western Michigan University; WILBUR B. BROOKOVER, Michigan State University

The independent variable of this study was school racial composition categorized as predominantly Negro, racial-mixed, and predominantly white. The indices of teacher influence, the dependent variable, were the magnitudes of associations between: (1) teachers' ability evaluations of students and students' self-conceptions of ability; (2) teachers' achievement preferences for students and students' achievement aspirations; (3) teachers' achievement preferences for students and students' achievement levels; and (4) teachers' evaluations of students' ability and students' achievement level.

The major variables controlled for were: (1) socio-economic composition of schools; (2) students' socio-economic status; (3) measured intelligence; and (4) teachers' and students' racial identity. The population investigated included 260 Caucasian and 178 Negro 6th grade students and their Caucasian teachers in sixteen lower class schools of a large midwestern city.

Findings support the view that racial composition of the school has an effect on teacher influence independent of the socio-economic composition of the school or of the socio-economic status of the individual students. It was also concluded that racially-mixing a school does not weaken the magnitudes of teacher influence or the achievement levels of Negro or Caucasian students.

3:00 to 3:55 P.M. **SESSION 20.2** **DIVISION C**

CLASSROOM INTERACTION II

CHAIRMAN:

Louis M. Smith, Washington University

PAPERS:

Change in Behaviors of First Year Interns

RUSSELL A. HILL, Temple University, and DONALD L. MEDLEY, Educational Testing Service

The study to be described was part of a larger project designed to try out a new instrument for observing and recording teacher behavior in the classroom. The subjects were seventy beginning teachers in a graduate teacher training program who were observed on the job during the second semester of their first year of teaching, twice in February and twice in May. This paper reports the results of an analysis designed to yield information about the behaviors of first-year secondary school teachers, changes in their behaviors over the semester, and relationships between teacher behavior and subject-matter being taught.

Two instruments were used: Flanders' Interaction Analysis Technique (FIAT) and Observation Schedule and Record 4 Verbal (OScAR 4V), one by each of two observers who visited the same teacher at the same time. OScAR was scored on 42 scales, FIAT on 44. Statistically significant overall changes were observed on 15 OScaAR and 14 FIAT scales. Subject-matter differences were found on one OScAR scale and 8 FIAT scales. Thirty-seven OScAR scales and 38 FIAT scales discriminated reliably between teachers of the same subject and grade level, and 19 OScAR and 25 FIAT scales provided reliable measures of individual differences in changes over the semester.

The resulting quantitative patterns of behavior are compared with popularly held concepts of teacher behavior. Speculations are offered about the common sense significance of several new keys in both the OScAR 4V and the FIAT found to be reliable in discriminating between teachers and between subjects. Finally, the direction of change indicated by the differences between the observations at the beginning and the observations at the end of the semester are examined for hypotheses about the development of teacher behavior during early career experience.

◆ ◆ ◆

Teacher Behavior and Pupil Achievement

EVAN R. POWELL, The University of Georgia

The purpose of the study was to examine the effects of indirect or direct teacher influence on pupil achievement. It was hypothesized that, for pupils under the same teacher for their first through third grade school experience, indirect teacher influence would be positively related to higher pupil achievement than would be direct teacher influence.

The teacher influence styles were determined through the use of a multivariate composite score derived from half-day observations using Flanders' Interaction Analysis. Teachers for the first three years, and for a subsequent fourth grade year, were dichotomized as indirect or direct, yielding four groups of pupils; Group I-D: three years of indirect teacher, one year of direct teacher, and Groups D-I, I-I, and D-D.

SRA Achievement tests were used to measure the dependent variable; Kuhlmann-Finch IQ tests served as a control variable. Groups were equalized at N=21 of each sex with a total pupil N=168. The achievement tests were administered after three years and again after a fourth year.

Using ANOV, groups were compared, with significant results yielding support for the supposition that three years of indirect influence under one primary teacher was associated with higher pupil achievement than three years of direct teaching. For the one subsequent year with a different teacher (in 4th grade), the results were mixed.

The conclusion reached was that while the long-term data lent support to the hypothesis in the direction of much other research, the fourth year (short-term) data were so mixed as to make any overall conclusion dubious.

◆ ◆ ◆

The use of Interaction Analyses in Student Teaching

EDMUND J. AMIDON, Temple University

This study tested the effects of training in interaction analyses upon the behavior and attitudes of student teachers in a secondary education program. The study determined whether student teachers who had been taught interaction analysis were perceived by pupils to be more indirect at the end of student teaching, to have more positive attitudes toward student teaching, and to be more indirect in their teaching than student teachers not taught interaction analysis. The study also determined whether student teachers working with cooperating teachers trained in interaction analysis differed on the above variable from student teachers working with cooperating teachers not trained in interaction analysis.

Forty secondary education student teachers were randomly assigned to one of four experimental groups during each of three semesters. Group I learned interaction analysis and was supervised by cooperating teachers trained in interaction analysis. Group II learned interaction analysis and was supervised by cooperating teachers trained in learning theory but not trained in interaction analysis. Group III did not learn interaction analysis and was supervised by teachers trained in interaction analysis. Group IV did not learn interaction analysis and was supervised by cooperating teachers trained in learning theory but not trained in interaction analysis. Student teacher behavior was assessed by observers using interaction analysis, and attitudes were assessed by the Teaching Situation Reaction Test. Student teacher personality was measured by Rokeach's Dogmatism Scale, and all students were rated on a rating scale. The Student Perception of Teacher Influence Scale was used to assess children's perception of student teacher behavior.

Of all measures used, the only one which identified significant differences between any of the groups was the interaction analysis. Those student teachers who had interaction analysis were observed as being indirect at the end of student teaching, as compared with those who had been trained in learning theory.

◆ ◆ ◆

The Attitudes and Teaching Patterns of Foreign Language Student Teachers Trained and not Trained in Interaction Analysis

GERTRUDE MOSKOWITZ, Temple University (Richard C. Anderson, University of Illinois)

The purpose of this study was to compare the attitudes and teaching patterns of a group of foreign language student teachers who were trained in interaction analysis with a group not so trained. Attitudes examined in the study were: (1) the student teachers' attitudes toward teaching and (2) toward their cooperating teachers; (3) the cooperating teachers' attitudes toward these student teachers; (4) the attitudes of the foreign language pupils toward the student teachers; and (5) foreign language as taught by them.

This study, conducted at Temple University, was carried out in the Fall semester, 1966, and replicated in the Spring semester, 1967. The subjects were all the foreign language pre-service teachers taking student teaching in either semester (N= 32). Each semester, an experimental group took a course in which they received training in interaction analysis, while a control group did not receive this training. After the experimental groups had received the training in interaction analysis, reliable observers who were trained in both foreign languages and interaction analysis, using the latter, coded two live classroom lessons of each student teacher: a grammar and a conversation lesson. In addition, appropriate questionnaires were administered to all subjects and their cooperating teachers and pupils to assess the aforementioned attitudes.

At the end of the training in interaction analysis, when compared with the control group, the experimental group:

1. taught both grammar and conversation lessons by using a variety of patterns which were more indirect;

2. had more negative attitudes toward their cooperating teachers, who did not know interaction analysis;

3. had pupils who perceived them and foreign language study more positively.

It appears as though training the foreign language student teachers in interaction analysis was related to their developing:

1. more positive teaching patterns;

2. more negative attitudes toward their cooperating teachers;

3. more positive attitudes in the pupils in their classes.

The results of this study are consistent with those of related studies conducted in other academic disciplines.

3:00 to 3:55 P.M. SESSION 20.3 DIVISION D

STUDIES OF LEARNING ENVIRONMENTS

CHAIRMAN:

Erwin H. Brinkmann, Southern Illinois University

PAPERS:

The Structure of Behavioral Maturity at the Pre-School Level

HARRY E. ANDERSON, JR., W. L. BASHAW, CHARLES E. JOHNSON, and WARREN G. FINDLEY, University of Georgia

Behavioral maturity of pre-primary children was studied with Kim's Behavioral Maturity Scale (BMS). BMS is a teacher rating device developed with second graders for evaluating academic (A), emotional (E), and social (S) maturity. These 3 factors were quite clear at the second grade and were more highly related to academic criteria than standardized tests. Each factor is measured by 6 items rated on a 5-point scale.

BMS was completed for all students in an experimental pre-primary school in which the sample of students was stratified on sex, intelligence, socio-economic status, and race. There were 58 3-year olds, 56 4-year olds, and 56 5-year olds. Data were analyzed by age group.

An examination of the principle axis roots revealed four roots above one in all analyses. Graphs (scree test) also supported a 4-factor solution. Both 3- and 4- factor rotations were conducted. Communalities below .6 were rare; most were between .7 and .9.

The 3-factor rotations yielded a clear simple structure of A-, E-, and S-factors for 4- and 5- year olds. For 3-year olds, the A-factor collapsed into the E- and S-factors. The third factor combined work completion items with items dealing with appropriate affect toward others.

The 4-factor rotations clearly yielded E- and S-factors for all three groups. In the 3-year old data, some A-items loaded high on factor E. Factor 3 was academic independence and book enjoyment while factor 4 was task completion. (These together compose the item pool for factor A.)

The 4-year olds yielded clear A-, E-, and S-factors. Factor 4 combined book enjoyment with appropriate affect toward others.

In 5-year olds, the A-factor split into book enjoyment (factor 3) and academic independence and task completion (factor 4).

Results are discussed in terms of similarities to the following samples of 8-year olds: American Caucasians, American Negroes, Koreans, and Japanese.

Generalizability of Performance by Job Corps Trainees on a Universe—defined System of Achievement Tests in Elementary Mathematical Calculation.

WELLS HIVELY II, Minnesota National Laboratory

Despite considerable recent pressure from activity in the fields of programmed instruction, computer-based instruction, and curriculum evaluation, the methodology of achievement testing still has not been brought to bear effectively upon the problem of behaviorial definition of instructional goals. Appropriate mathematical models have existed for some time, based on the notion of domain sampling and on the analysis-of-variance approach to test reliability, or "generalizability" (Buros, 1963; Cronbach, Rajaratnam, and Gleser, 1963; Lindquist, 1953). What is needed is for behaviorists to take these models seriously.

If one thinks of an achievement test item as a sample from a general class of behavior, he may proceed to write the rules which define the set of all equivalent items taken to represent that class. Call this set of rules a test-item form. An area of subject matter may be operationally defined as a collection of item forms. These in turn may provide the sampling domain for a system of universe-defined achieve-tests (Hively, 1966; Osburn, 1967).

This paper presents data from a generalizability study (Cronbach, Rajaratnam, and Gleser, 1963) of a system of tests of elementary mathematical calculation based on the above approach and designed for use in the Basic Education Program of the Job Corps' Youth Conservation Centers.

Item-form sampling domains were defined for tests in each of the subject-matter areas listed in the section below. Three random-parallel test forms were generated from each of these domains and used as repeated measures in generalizability studies on samples of Job Corps trainees ranging in number from 54 to 112.

A rigorous, domain-sampling approach to the construction of achievement tests in elementary mathematical calculation was found to be feasible, yielding reliability coefficients between .82 and .90, each of which permits generalization to performance on an essentially infinite family of randomly generated, parallel, test forms. Practical and theoretical implications are discussed.

◆ ◆ ◆

Playfulness and Nonplayfulness in High-school Students: Trait Composition and Educational Implications

J. NINA LIEBERMAN, Brooklyn College of the City University of New York

The importance of identifying behavior which provides a clue to cognitive style, especially divergent thinking, was the rationale for investigating the manifestations of playfulness-nonplayfulness in the HS classroom.

The present paper deals with the composition of pf-nonpf in HS students as well as the influence of sex, age, grade and subject matter area on this behavior.

Three-hundred and thirty-eight boys and 272 girls from grades nine through twelve, ranging in age from 13.1 to 19.3 and representing four achievement levels, were rated twice by their classroom teachers on a specially constructed instrument, covering ten traits of pf-nonpf and two unrelated questions. Before the second rating session, the teachers were also asked to rank their students according to their conceptualization of pf-nonpf.

A principal components factor analysis with Varimax rotation to four factors was performed on the test scores, age, sex and grade of student, and sex of teacher. Two distinct and psychologically meaningful factors emerged for pf-nonpf. The first factor was labeled social-emotional pf-nonpf and the second, academic pf-nonpf. The latter was saturated also with the ringer question on achievement-orientation.

No significant correlations were obtained between pf-nonpf and the sex, age and grade of student, or sex of teacher.

Two one-way analyses of variance assessed the influence of subject matter area and type of school on the pf-nonpf ratings. Except for physical alertness and enthusiasm, differences among subject matter areas were significant at the .01 level on all scales. No significant differences were found between urban and suburban type of school. Test-retest reliability was .82 for total test, with item reliabilities ranging from .36 to .86. Interrater reliability on a subsample of 158 showed an item mean of only .38. Validity coefficients obtained between total test and re-test scores and teacher-rankings were .69 and .76, respectively.

The educational implications of a two-factor syndrome of pf-nonpf in HS students point to situation-specific (academic) vs. situation-spanning (social-emotional) traits. In the follow-up study the relationship of these two clusters to divergent thinking will be explored. Also of importance is the influence of subject matter on pf-nonpf, suggesting that subjects differ in the amount of "play" allowed in approaching them.

◆ ◆ ◆

A Longitudinal Study of the Factorial Composition of Intelligence in Elementary School Children

LANCE W. HEMBERGER, University of Georgia (R. Travis Osborne) of Georgia

For the purpose of identifying qualitative differences in the dimensionality of the intelligence test performance of elementary school children as a function of increasing educational experience, 163 white children were selected randomly from small rural, medium and large urban communities. These children were first tested during the summer prior to entrance into the first grade. All available members of the original group were examined again at the completion of the first, third and fifth grades.

While the Wechsler Intelligence Scale for Children (WISC) was the primary instrument employed, the Peabody Picture Vocabulary Test and the Ammons Full-Range Picture Vocabulary Test served in the battery as reference tests.

Prior to statistical analyses, each WISC subtest (except Coding)was split into two, three or four experimentally independent parts to which individual items were assigned in simple rotation. These additional subtest splits provide the design with a sufficient number of variables to establish at least as many factorial dimensions as there are subtests, but does not force this result.

The test data for each grade level were factor analyzed and rotated to Kaiser's varimax criteria. Tucker's method of maximal congruence was employed to facilitate comparison of the resultant matrices.

Results of the present study were generally compatible with previous WISC factor studies. Eight statistically significant orthogonal factors were extracted from the pre-school data, ten factors from the first grade, nine factors from the third grade, and eight factors from the fifth grade data. Evidence is presented supporting the existence of stable mental factors from the pre-school level through grade five. Results indicated greater factorial complexity than has frequently been imputed to the Wechsler, lending support for the practice of clinical interpretation of the Wechsler psychogram.

3:00 to 3:55 P.M. **SESSION 20.4** **DIVISION B**

CURRICULUM EVALUATION

CHAIRMAN:

Leonard S. Cahen, Educational Testing Service

PAPERS:

An Evaluation of the Continuous Progress Concept of Instruction with University Students

J. WILLIAM MOORE, J. MAURICE MAHAN, C. ALLEN RITTS*,
Bucknell University

The purpose of the investigation was to determine the effects of the Continuous Progress Concept of instruction on University students' achievement and attitudes.

The experimental program which arranged course content along a continuum required students to master the concepts as they progressed through the program prior to moving on to concepts at the next level of difficulty. To provide for individual styles of learning and learning rates, the instructional procedures were developed in an individualized instructional format.

Three courses in each of three departments, Biology, Philosophy and Psychology, were involved in the experiment. Three courses were selected for programming in the continuous progress format to afford students the opportunity to go beyond the boundaries of a single course if they completed (because of the self-pacing dimensions of the program this was possible) the first course before the end of the first semester.

Experimental and control groups for Biology and Psychology were assigned using a table of random numbers, and they participated under the respective experimental and control conditions during the same period of time. Because of scheduling difficulties, the reference group for the experimental program in Philosophy could not be determined by random assignment. For this reason it was necessary to use the academic performance of Philosophy students completing the comparable course during the previous semester as a basis for comparison.

The following results are the findings of the first semester of the experimental program.

Post-test (final test) scores for experimental and control groups were compared statistically for both the Biology and Psychology programs. In both instances the mean achievement scores for the students completing the continuous progress courses were significantly higher than the mean achievement scores of students in the respective traditional programs. Similar findings were observed in the Philosophy program with the mean achievement of the continuous progress students being higher than the mean achievement of the students who completed the course under traditional conditions.

In an attempt to gain further evidence of the effects of the experimental conditions in the academic performance of students, achievement scores on unit tests were compared for Biology and Psychology students with similar results.

Further, the experimental groups tended to have the most favorable attitude toward the continuous progress approach to instruction; however, both groups, i.e., experimental and control, strongly favored the continuous progress approach.

In summary, it can be concluded that the requirements of mastery of course objectives coupled with a program of individualized instruction, result in an increase in academic performance and a more favorable attitude toward the instructional process.

◆ ◆ ◆

The Effect of the High School Curriculum Upon College Achievement

ROCK ASHCRAFT, New Mexico State University

The principal objective of this study was to evaluate the effect of two different high school curricula, college preparatory and non-college preparatory, upon college achievement. Further, the correlation of these two curricula was investigated with relation to the particular college of the university in which the students were enrolled. Separate male and female analyses were conducted to determine whether the factor of sex significantly effected achievement between groups and among colleges.

A complete record of the high school transcript and ACT scores for each of the 906 subjects was secured from the office of the college registrar. Each of these was examined, and the amount of time spent by the student in the different courses was converted to standard Carnegie units. Each transcript was then evaluated, and the results tallied as to the number of Carnegie units in the college preparatory area and the number in the non-college preparatory area. This evaluation included four years of high school work from the ninth through the twelfth grade. The total units of high school credit divided into the number of college preparatory units provided a measure of the high school curriculum background for the student, expressed as a percentage figure. All courses attended by the student were included.

The records were then arranged in rank order, according to the percentage of college preparatory content included in the student's high school curriculum. Divisions were made at the first and third quartiles according to acceptable standards for establishing groups from ordinal data. This division resulted in the establishment of Group A as those students completing *73 percent or more* of their high school work in the college preparatory area, and of Group B as those students completing *59 percent or less* of their high school curriculum in the college preparatory area. All students having the same percentage of college preparatory units as the student who was actually on the first and third quartile points were included in the sample. The number of subjects in the study was then established as 228 in Group A, and 226 in Group B.

Analyses of covariance of college Grade Point Average were conducted on the combined male and female sample, and on separate male only and female only samples for eight consecutive semesters of college. The American College Test composite score was employed as the concomitant variable for statistical control of ability in the analysis of covariance. Various supplementary analyses of the basic data were also computed on an a priori basis.

The two groups were highly different in their ability to perform college tasks as measured by the American College Test with Group A excelling over Group B. When the difference in ability was adjusted to a common mean in an analysis of covariance, the achievement of the two Groups was not significantly different. The Group by college interaction values were not significant in the analyses of variance or the analyses of covariance, indicating that achievement was not dependent upon a certain combination of group and college. When ability, as indicated by the American College Test, was adjusted to a common mean among colleges, no significant differences in grade point average was attained by the student because of his selection of a particular college. However, isolated exceptions to the above conclusion were found.

The separate male only and female only samples appeared to be from the same population as the combined male-female sample.

Analyses of variance of college grade point average were conducted for the male only and the female only subsamples. The results indicated that the first year of college was the most critical concerning the relationship of achievement to high school background. Highly significant differences in ability existed between the two groups; however, achievement was not significantly different for the remaining six semesters.

In summary, the data do not tend to indicate that the high school background is highly critical to college success, but tend to indicate that general intelligence and non-intellective factors are more decisive to college achievement.

Recommendations pointed to the need for additional research in the area of high school background as it may affect the college dropout rate.

Performance on Modern Language Association Tests by Regular and FLES Students in French

JOHN M. NEWELL, University of Florida, Gainesville, and EVELYN BREGA, Lexington Public Schools, Lexington, Mass.

A previous paper by the authors presented results indicating the superior performance of FLES students on MLA tests. In that paper, not only was the number of students tested small (15 and 17) but differences in mean IQ for the two groups were statistically different.

The present study is concerned with two groups of 54 students (21 males and 33 females). One group began French in grade nine and is presently completing its third year of French. The other group (FLES) was introduced to French in grade three and has continued to the present. The problem in this study is to determine if exposure to FLES programs results in long-term superior performance on MLA tests when difference in IQ of subjects is statistically controlled.

An analysis of covariance for each of the four MLA tests (Listening, Speaking, Reading and Writing) was performed with IQ as the concomitant variable. Results support the findings of the previous study in that FLES students performed significantly better on all four MLA tests. With differences in IQ statistically controlled, the present findings provide useful and needed research evidence of the superiority of FLES students on MLA tests. The results also offer evidence for the long-term effect of FLES on performance on the MLA tests. No significant differences for sex were found. One FLES class and one regular class was taught by the same instructor. Significant differences in performance between these two sub-groups were found although, for each sub-group, no significant differences in scores were found when each sub-group was compared to the total group (FLES or regular). This suggests that instructor differences were not the critical factor in this study.

A description of the FLES program and methods of data collection are presented with suggestions for modifying existing audio-lingual programs to give greater emphasis to non-verbal skills.

◆ ◆ ◆

Pupil Achievement Attitude and Behaviors in a Multi-Age Nongraded School

FRANCIS X. VOGEL, Florida State University, Tallahassee, and NORMAN D. BOWERS, Northwestern University

In speculative writings it has been argued that, as compared with the traditional form of school organization, the nongraded form should be superior in developing pupil classroom behaviors that are related to generally accepted educational objectives. The idea crucial to the nongraded form is that multi-age classroom groups occur for each level of achievement. As is well known, in the traditional graded form of school organization, the pupils' ages are held relatively constant, and the achievement levels within any classroom group show wide variability. This study sought to determine the relationship of these forms of school organization to selected classroom behaviors of pupils.

Specifically, multivariate analyses of covariance were used to analyze data that compared overage, normal age, and underage pupils in non-graded classrooms with overage, normal age, and underage pupils in traditional classrooms. The comparisons utilized scores on measures of conceptual maturity, attitudes to school, achievement, group participation, and responses in actual classroom situations that were classified as contributing or noncontributing. The specific measures were the *Describe Your School, Stanford Achievement Test, The Draw-A-Person Test, The Russell Sage Social Relations Test* and two scales derived from data on the *Observation Schedule and Record (2-e).* Data were collected at the beginning of the school year on the *Describe Your School, Stanford Achievement Test,* and *Draw-A-Person Test.* These were used as covariates; data collected subsequently during the winter and spring were used as dependent variables. The multivariate analysis of covariance indicated significant interaction of age groups with school groups ($P<.05$), differences among the age groups ($P<.01$), and differences between the school groups ($P<.01$). On examination the data indicated that the two classroom observations scales were the only measures which contributed to the significant interaction. When a second multivariate analysis was performed that eliminated these two measures, no significant interaction was found, and significant differences were obtained for age groups and school groups ($P<.01$). Additional univariate analyses were completed that indicated directionality of differences and aided in specific interpretations. The nongraded form of organization appeared to encourage pupil development in conceptual maturity and participation in group activities. The graded form of organization seemed to encourage pupil development in the more traditional measures: achievement, attitudes toward school, and the scale involving teacher directed activities during the usual classroom operation. The data relating to differences among age groups can offer some argument that grouping pupils on a multiage basis in each classroom would be desirable. These findings offer some quantitative evidence relating to important issues of validity of the nongraded school idea.

◆ ◆ ◆

An Exploratory Study of the Relationship of Belief Systems, Goals, and the Evaluation of College Undergraduates

JAMES A. CONWAY, State University of New York at Buffalo and JOHN R. DETTRE, The University of New Mexico

Ebel, Alexander, Getzels and Jackson, Battle, Rokeach, and others have suggested that dogmatism or "open and closed mindedness" may be directly involved in the biases and idiosyncrasies that serve to influence marks given as a result of teacher-pupil relationships in the learning situation.

The purpose of the study is to determine the extent to which congruence of belief systems of students and of teachers (or the lack of congruence) interacting with similarity of goal orientation of the same students and teachers (or the lack of similarity), influences the final mark given in courses in teacher education on the undergraduate level.

Twenty-six faculty and 792 students at the State University College at Buffalo, New York, were administered Rokeach's Dogmatism Scale and an adaptation of the life goals used by Getzels and Jackson. Midway through the semester faculty Ss were asked to complete another questionnaire. They were asked to rate each of their students as being above average, average, or below average; ratings were to be general impressions of the student without reference to grade books or other sources. At the end of the semester final grades were collected for each student subject.

Extremes on the D Scale and extremes in terms of goal agreement or disagreement were selected for analysis. Combinations of Closed Faculty, Open Faculty, Closed Students, and Open Students were formed for chi square analyses of grade distributions. High agreement and Low agreement groups were likewise formed as well as groupings stemming from the interaction of belief systems with goal agreement as related to final grades. Similar groups were also used for looking at faculty estimates of students prior to final grades. Distributions of grades among the groups formed by the interaction of belief systems and goals in relation to grades were significantly different (.05 level). Mid-semester estimates of groups formed on the basis of open and closed mindedness also resulted in significant differences (.01 level). In general it appeared that closed minded subjects received higher evaluations than open minded Ss. The greatest differences occurred in favor of those students who were closed minded *and* in low agreement with the goals of their instructors.

The results lead to some speculative argument. Could it be that the closed minded students, in low agreement with their instructor's goals, were quite verbal in the class situation? Further, could it be that these vocal students become quite visible through class participation and as such, are looked upon with favor by their individual teachers? The results, together with other information, lend some support to this argument.

3:00 to 3:55 P.M. SESSION 20.5 DIVISION D

CHARACTERISTICS OF STUDENTS

CHAIRMAN:

Frank Costin, University of Illinois

PAPERS:

Assessing Student Perception of the College Experience: A Factorial Approach

S. THOMAS FRIEDMAN, BERNARD HOROWITZ and DAVID SHERRILL, The University of Texas

In the effort to obtain greater insight into student perception of the college experience the College Attitude Survey (CAS), a 165 item attitudinal questionnaire was constructed and administered to a relatively representative sample of undergraduates at The University of Texas (N= 739) in the spring of 1967.

The 165 item CAS, with each item requiring a response on a 5 point bipolar scale, consisted of items the authors considered relevant to the major dimensions of the college experience.

All correlation coefficients between all pairs of the 165 items were computed. The correlation matrix thus obtained was subjected to factor extraction by the principal axis method and those factors having eigenvalues of 1.00 or higher were rotated by Kaiser's varimax method.

The 43 common factors were studied and tentatively identified. Scale scores were then obtained for each student accross the 43 factors using an equal weighting procedure involving all items whose loading on a given factor exceeded ±.30.

The set of 43 scale scores were then subjected to a second order factor analysis yielding the final 13 factors of the instrument. The second order analysis was obtained also by the principal axis method with those factors having eigenvalues of 1.00 or higher rotated by Kaiser's varimax method.

Among the 13 factorially derived scales are: (1) student perception of faculty-student interaction, (2) acceptance of the establishment, (3) religious activity, (4) moral skepticism, and (5) press for academic success.

First and second order scales, together with reliabilities, are presented. Differences between various sub-sets of the sample are also indicated.

◆ ◆ ◆

Toward Objective Student Typologies: A New Kind of Configuration Analysis

RICHARD O. GOODWILLIE, University of Connecticut

In recent years several authorities have pointed up the methodological promise of configuration analysis, suggesting that we may be able to learn more about the real-life interaction of social and psychological variables if we begin by distinguishing among different overall configurations of such variables—i.e., if we first employ clustering techniques to identify relatively homogeneous subgroups, or types, within the population. Biographical information, test scores, and other student data are likely to have more meaning and more predictive power when analyzed for such subgroups than when based on an undifferentiated group. Yet, perhaps because of limitations in existing methods of configuration analysis, very little research of this kind has been attempted.

This paper reports the development and initial application of a new form of configuration analysis designed to yield meaningful typologies. It is based upon Cattell's concept of types as modal configurations of traits—i.e., patterns occurring more frequently than other possible patterns. A matrix of similarity indices is produced, one for each pair of persons in the population, each index, reflecting the *distinctiveness* as well as the *number* of shared characteristics between two persons. Special clustering procedures are then applied to the matrix, and the resulting clusters are studied for evidence of differentiating patterns.

Biographical and psychometric data (over 100 variables) were analyzed for two groups of engineering freshmen, 233 successful students and 221 unsuccessful students, yielding several types in each group. Although validation awaits further investigations, the results provide evidence of the method's feasibility, of its ability to disclose psychologically meaningful patterns, and of non-linearity in the operation of some key variables. Possible implications for further research and for educational practice are discussed.

◆ ◆ ◆

The Perceptions of School Personnel and Parents Toward Behavior Problems of Children as Compared to Mental Hygienists

GERALD G. MANSERGH, The University of Michigan

This study was developed to assist in the implementation and evaluation of an exemplary school mental health program involving the use of mental hygienists in helping school personnel and parents develop more positive attitudes toward children.

A modified version of the Wickman Behavior Rating Scale was used to obtain the perceptions of school personnel, parents, and mental hygienists toward the importance of certain "inferred" behaviors in the future adjustment of elementary and/or secondary age children.

Within a five school district area of suburban Minneapolis, the instrument was administered to all elementary principals (63), all elementary school psychologists (10), a random sample of elementary teachers (137), a random sample of parents of elementary children (111), all secondary principals (19), all secondary assistant principals (24), all secondary counselors (76), a random sample of secondary teachers (127), and a random sample of parents of secondary youth (71).

In addition, a mental hygienist criterion group of psychiatrists (10), clinical psychologists (6), and psychiatric social workers (22) from the six major mental health clinics in the Twin Cities metropolitan area were administered both the elementary and secondary forms of the instrument.

Utilizing a multiple linear regression computer program, "inferred" behavior ratings of the various groups of school personnel and parents at the elementary level were compared one to another and to the criterion group of mental hygienists; similar comparisons were made with the rating of secondary level respondents. A level of .05 was considered a significant difference.

General findings were as follows:

1. School psychologists and counselors were closest to mental hygienists in their perceptions; all were more concerned with recessive behaviors.

2. An extreme number of significant differences exist between the perceptions of elementary principals or parents and school psychologists. The former groups are significantly more concerned with aggressive behaviors.

3. At the secondary level, mental hygienists and parents or teachers showed the greatest numbers of significant perceptual differences.

4. Among secondary respondents, the greatest number of significant differences in ratings was found with the counselor-parent comparison.

◆ ◆ ◆

Interaction Analysis of Mexican-American and Negro High School Students to College Orientation

ROBERT WARD, University of California, Santa Barbara

A group of highly able non-college bound Mexican-American and Negro high school students were brought to the University of California, Santa Barbara for a two day college orientation. The group of thirty-eight high school students was broken into small groups headed by college counselors pre-trained in the use of simulated college orientation "games." Tapes were taken of all meetings and an interaction analysis was done using Bion's categories and Flanders' Verbal Interaction Categories. The findings suggested interesting emotional reactions of these minority-group youngsters toward college. Even more revealing was the effect various counselor systems had on group interaction.

3:00 to 3:55 P.M. **SESSION 20.6** **DIVISION E**

COLLEGE STUDENT ADJUSTMENT AND DEVELOPMENT

CHAIRMAN:

Edward F. Krahmer, University of North Dakota.

PAPERS:

Student Persistence in Voluntary Learning Activities: A Study in Hierarchical Value Structure

SEYMOUR METZNER, City University of New York

During the past five years, "free universities" or non-credit, non-institutionally sponsored courses, have appeared on several campuses. After an initially enthusiastic response, they have tended to lose student interest. This study was undertaken to cast light on possible relevant factors involved in the loss of student participation.

An analysis was made of nonreactive factors which were hypothesized as bearing upon college student persistence in 29 voluntary, non-credit, after-school classes initiated by a student government with volunteer faculty participation. The factors investigated were the day of week and time of day the classes were held, amount of assigned readings, and content of the courses.

All the hypothesized factors were found to be significantly related to student persistence, but did not sufficiently account for the trends in student attendance. A subjective, value oriented factor, as delineated in personal interviews, seems to have been more important. This was particularly noticeable in the effect on student persistence of other curricular and non-curricular activity demands.

◆ ◆ ◆

Peer Integration and Academic Success: Toward a Sociological Model of Undergraduate Attrition

WILLIAM G. SPADY, Harvard University

Despite extensive screening and highly selective academic standards, the college of a renowned university loses at least ten per cent of its freshmen (plus many older students) annually. A theoretical model was developed which treated the college as a social system with distinctive values and expectations into which newcomers would need to integrate.

The model consisted of five variables, each hypothesized to exert a positive independent influence on remaining in the college beyond the first year. These variables included: *grade performance,* the student's subjective sense of *intellectual development;* the compatability of student values with dominant values and characteristics of the college environment *(normative congruence);* the extent and intensity of interaction with "close friends" *(friendship support);* and the student's *perceived* sense of *integration* in the college social structure. Academic potential (a composite of high school class rank, SAT scores, and a measure of high school quality) was used as a control.

Data were gathered on all 683 entering freshmen in September, 1965. Information from two questionnaires administered before entrance; and during April of the freshman year (97 percent returns on both) was augmented by complete admissions records and college grades. Except for normative congruence and grades, the variables were constructed by principal component analysis of theoretically homogeneous sets of items, and were analyzed in both continuous (regression analysis) and ordinal form (percentage analysis).

The absolute and relative effect of each variable in the model differed between men and women, although each was positively related to staying in school for both sexes with the effects of the remaining variables held constant. Although males were more affected by grade performance and intellectual development, females were particularly influenced by normative and interpersonal supports.

Analysis of major field groups by sex revealed differences which could be interpreted according to the concept of relative deprivation. The higher the expressed "achievement motivation" mean scores of male groups, the more likely were low grade performers and intellectually unfulfilled students to leave. Similarly, for both sexes, the greater the "interpersonal orientation" mean scores of groups, the more likely were those with low friendship support and low perceived integration to drop out.

The utility of examining the college as a social system with potential for rewards and supports in a variety of cognitive and interpersonal activities is apparent.

◆ ◆ ◆

A Theoretical, Practical Value Characteristics Inventory: A Factor Analytic Study

R. J. PLANISEK and P. R. MERRIFIELD, Kent State University

The purpose of this study was to provide information concerning the scholastic, vocational orientation characteristics of college students. The existence of an attitude or value structure in theoretical or practical domains was sought. In previous studies, personality measures which have a significant correlation with measures of academic performance often reveal a theoretical or practical item content. If students may be described on a continuum from "intellectually curious" to "vocationally inspired," then measures of a student's theoretical, practical value structure may have academic correlates. An instrument of 40 items was constructed to assess these theoretical, practical value characteristics.

The sample of 103 Kent State University men attending summer school consisted of equal numbers of freshmen and upperclassmen. These subjects evaluated each value statement on a 9-point scale indicating their degree of disagreement, ... , no opinion, ... , agreement. Means, standard deviations, and frequency distributions of responses were computed for each of the 40 statements; then product moment correlations were computed between each pair of the 40 items. The matrix was factor analyzed by the principal components method using unity in the diagonal. All factors with eigenvalues greater than 1.00 were rotated to the varimax criterion.

Upperclassmen obtained higher means on theoretical and lower means on practical (item) scale scores than did the freshmen; however, these were not significant at the .05 level. The 8-factor solution was interpreted as follows: Intellectual-Research, Intellectual-Historic, Theoretical-Academic, Intellectual-Communication, Vocational-Research, Technician-Practical, Practical-Academic, Technological-Societal.

The existence of these value characteristics was supported as separate but independent characteristics by the degree of simple structure in the orthogonal solution. However, an oblique solution may produce theoretical, practical second-order factors which display some degree of polarity.

Presently, studies in profile distributions and academic correlates are under investigation.

◆ ◆ ◆

The Effect of Correlated Class Scheduling and Dormitory Housing on the Social and Academic Behavior of College Freshmen

ROBERT G. STAKENAS, Florida State University

Currently it is assumed that large university size impedes the intellectual growth and personal development of students. In light of this problem, the purpose of this study was to determine the effect of an organizational procedure termed *clustering* upon the social and academic experiences of incoming freshmen at a large university.

Following Newcomb's theory of peer group formation, students were clustered into groups of 25-30 with four courses in common. The purpose of clustering was to maximize opportunities for social interaction and to provide common academic experiences. To maximize opportunity for interaction, four clusters were housed in a co-educational living-learning dormitory. Other cluster students were randomly assigned to dormitories scattered about the campus. Appropriate control groups were also selected.

The study sample was surveyed at the close of both the first and second semesters by means of paper and pencil questionnaires. The obtained data were analyzed by the analysis of variance.

Systematic differences between groups failed to emerge for the outcome measures of perceived stress, informal social participation, extracurricular involvement, or grade point average. However, results based on measures of peer pressure and general satisfaction tended to confirm expectations derived from Newcomb's peer group theory. Level of peer pressure was related to opportunity for social interaction, and tended to increase over time. Living-learning clusters reported highest general satisfaction the first term, and controls the lowest. Differences in satisfaction tended to diminish during the second term, although cluster students remained most satisfied. Initial differences in satisfaction might be due to a possible Hawthorne effect.

It is tentatively concluded that the organizational procedure of clustering provides the opportunity for enhancing general satisfaction in college, and can also lead to higher degrees of peer influence. However, general satisfaction and peer influences do not appear to have any significant effect upon the academic behavior of students.

3:00 to 3:55 P.M. **SESSION 20.7** **DIVISION D**

STUDIES OF NON-AMERICAN SUBJECTS

CHAIRMAN:

James R. Bower, University of Michigan

PAPERS:

Transposed Factor Analysis as a Means of Typing Educational Philosophies

WILSON H. GUERTIN, University of Florida

Dissertation data have shown the reactions of Jordanian and Florida student teachers to educational philosophy statements. Analysis failed to reveal large item intercorrelations that could be attributed to personal beliefs and attitudes such as stinginess, progressiveness, and democracy. The failure to find such dimensions was attributed to the centuries of powerful social-cultural influences that have been over-riding individual traits to produce a uniform cultural product.

By intercorrelating individuals in a small sample of Arabs and Americans instead of items, one would expect Arabs to look more like one another than like Americans and vice versa. The tightness within each cluster reflects the amount of uniformity of philosophical belief within that culture. The looseness indicates freedom of individuals to feel differently from others within that culture. The overlap between clusters reflects the universality of educational philosophy across cultures.

A sample of 25 Arab and 25 American female student teachers gave high correlations among subjects from the same culture but correlations between members of the different national groups tended to be very low. The transposed factor analysis revealed two major independent factors. On the Arab type-factor all but six of the 25 Arab subjects loaded better than .50 and none of the Americans loaded higher than .47. The American type-factor showed all but five of the American subjects loading better than .50 while the highest Arab subject loading was .45.

It is apparent that for the sample used, there was: (a) distinct national educational philosophies; (b) no indication of a universal educational philosophy; and (c) no difference in the amount of individual deviation from the national pattern for either sample.

◆ ◆ ◆

A Comparative Study of Maturity in South Korean Second Grade Students

YUNGHO KIM, American Institute for Research, HARRY E. ANDERSON, JR. and W. L. BASHAW, University of Georgia

A Behavioral Maturity Scale (BMS), reflecting Academic (A), Social (S) and Emotional (E) maturity, was developed with more than 500 second grade American students and reported at the NCME conference last year. The scale has since been used at the pre-school level and with foreign students. The present study presents results of the BMS with 220 South Korean second grade students.

The BMS is composed of 18 items, six items each in the above three factors. The items were arranged in a random sequence and teachers in four classes were requested to rate their students on a seven-point scale for each item completing only five or six ratings a day, approximately at mid-year. All four second grade classes were in Kobuk Elementary School in Sosan Province, a fairly typical rural South Korean school.

The scale values for the items were intercorrelated and factor analyzed. Three- and four-factor solutions were obtained; three factors being hypothesized but four roots greater than one appeared in the analysis. Similar to other research with the BMS, high communalities were obtained for all items, ranging from .60 to .89, and the factors accounted for approximately 75 per cent of the total variance. Although a clear simple structure was obtained, the hypothesized three factors did not result in the analysis. One factor consisted of A and E items. A second factor represents S items. The third factor consisted of A and S items while the fourth factor was fairly specific for an E item. The results are compared with other studies including one with Japanese students, where the hypothesized factors were obtained.

◆ ◆ ◆

Evaluating Culture-Fairness in Translations of College-Level Reading Tests

TURKAN K. PEYSER, Illinois Institute of Technology

The purpose of the research was to study the comparability of an English college-level reading comprehension test to its Turkish translation administered in Turkey, and to its Enlgish translation administered in the U.S. This comparison would yield data on the relative culture-fairness of translations of reading tests for a better evaluation of the verbal ability of foreign students planning to study in the United States. Testing with two equivalent forms of reading tests, one in English, and one in the native language of the examinee, might provide a powerful diagnostic tool for identifying (a) the individual's potential for higher education, and (b) the extent to which this potential is depressed as the student is faced with the necessity of shifting from his native language to English as a language of instruction.

The instruments central to the study were two parallel forms of a reading test suitable for the level of college entrance in the United States. These tests were translated into Turkish, and re-translated from Turkish back into English. Supplementary measures employed were SAT-Verbal scores, intellegence scores, a vocabulary test score, a questionnaire, and school grades. The sample consisted of approximately 700 Turkish high school seniors, 200 Turkish college students, 400 American high school seniors, and 700 American college students.

Comparisons based on different psychometric criteria showed the following results: In terms of total test scores, difficulty of reading passages, and indices of item difficulty, the tests yielded relatively consistent results with translation and administration in a different culture. Responses to the wrong options of an item tended to be based on general knowledge more than responses to right options. The sharpness of discrimination of items showed a negligible relationship between cultures, although within each culture it showed considerable stability. Item difficulty and discrimination correlated low within the United States, but significantly within Turkey.

The findings may imply that careful translations of college-level reading tests measure an inferential ability that is relatively uniform throughout cultures.

◆ ◆ ◆

A Cross Cultural Study of Maturity

ALLYN PRICHARD, HARRY E. ANDERSON, JR., W. L. BASHAW,
University of Georgia

An 18-item child behavior scale, measuring Academic, Social and Emotional Maturity with six items, respectively, was developed by Kim, Anderson, and Bashaw, and a study of the maturity of second grade children was made.

The scale was translated into the Kanji Script, and was used in rating a group of 100 second graders in Takaoka, Japan. The three factors previously obtained in the American study at the second grade level were distinctly apparent.

The range of item factor loadings for the six appropriate items on each of the three emergent factors was as follows: Factor I, .76 to .84 on Academic Maturity and only .05 to .30 on the other two factors; Factor II, .71 to .80 on Social Maturity and only .01 to .33 on the other two factors; Factor III, .61 to .84 on Emotional Maturity and only .03 to .41 on the other two factors.

Correlations indicated a higher relationship between grades in each of seven academic areas and academic maturity scores than between grades in each of the academic areas and ability measures. Correlations also indicated a higher relationship between total grade scores and academic maturity scores (.762) than between total grade scores and ability measures (.481).

Canonical correlations indicated a good prediction of the seven school grades from the three factor maturity scores (.833).

The behavior scale appeared to be reliably applicable for use in cross cultural studies.

3:00 to 3:55 P.M. **SESSION 20.8** **DIVISION E**

TEACHER CHARACTERISTICS

CHAIRMAN:

Charles E. Skipper, Miami University

PAPERS:

An Exploratory Study of Relationships Between Students' Perceptions About School, Self, and Student-Teacher Relations in Grades Six Through Twelve

DON E. HAMACHEK and JAMES CONLEY, Michigan State University

A questionnaire designed to tap perceptions about self, teachers, and student-teacher relations was administered to 1300 students in grades 6 through 12 in an all-white middle class school system, in order to examine relationships which may exist between the way students feel about school and their school performance. An additional purpose was to assess whether or not there were differences in perceptions from grade 6 through 12. Results were analyzed according to grade, sex, and father's occupation.

Student perceptions become increasingly more negative from grade 6 through grade 12 in the following areas;

1. In terms of teaching skills, *teachers are perceived as less* a) helpful; b) stimulating; c) well-organized; d) enthusiastic; e) inspiring; f) able to communicate.

2. In terms of student-teacher relations, *teachers are perceived as less* a) human and friendly; b) understanding; c) interested in students.

3. 90 percent of all 6th graders like school (N = 201), but only 63 percent of all 12th graders do. (N = 101)

4. 43 percent of all 6th graders work as hard as they can in school, but only 15 percent of all 12th graders say they do.

5. 82 percent of all 6th graders do their homework regularly, but only 42 percent of all 12th graders do.

6. In no instance do students end up with *more* positive school-related perceptions than they started out with, except 11th and 12th grade girls who perceive teachers as *liking* them more than do girls in grades six through ten.

7. The higher the grade-point-average for either boys or girls, the more positive are perceptions related to school, teachers, and self.

Attitudes and Motivations of Teacher Candidates in Early Childhood, Elementary, and Secondary Programs: An Exploration

O. L. DAVIS, Jr., The University of Texas at Austin and KAORU YAMAMOTO, The University of Iowa

Studies of similarities and differences in attitudes and motivations among those preparing to teach at different school levels would seem to be useful in the analysis of existent programs and examination of the relationships between characteristics and teaching behavior of candidates from different preparation programs.

Ss in this study were 132 female undergraduates at a midwestern university. All Education juniors, [illegible] were early childhood majors, 40 were elementary education majors, and 61 were in the secondary education program. Administered to all Ss were: Bowers' *Teacher Opinion Inventory,* the *MTAI,* and Stern's *Teacher Preference Schedule,* Forms G and A.

Fourteen of the 22 test scales differentiated among the three groups. Early childhood candidates exceeded secondary candidates on satisfaction toward teaching, attitudes toward pupils and schoolwork, dominance, orderliness, dependence, exhibitionism, and preadult-fixed. On only the *MTAI* did early childhood candidates exceed elementary majors; on no scale was the reverse true. Elementary candidates scored higher than secondary candidates on *MTAI* and the Stern-identified motives of orderliness, dependence, exhibitionism, and preadult-fixed. Secondary candidates did not exceed elementary majors on any scale, but were higher than early childhood majors on two motives. These results are related to teacher education student personnel practices and curricular differentiation.

◆ ◆ ◆

Comparison Between Selected Characteristics and Performance of Provisionally and Professionally Certified Beginning Teachers in Georgia

JOSEPH C. BLEDSOE, University of Georgia

The purpose of the study was to compare selected personality characteristics and teaching performance of provisionally and professionally certified beginning teachers in Georgia.

A least-squares ANOVA model with 33 variables and a sample of 470 teachers was used. Main effects and first-order interactions were obtained for teacher sex, certification status, field of teaching (elementary, English, mathematics, science, and social studies), and place of training (Georgia or non-Georgia colleges).

23 differences (11 significant) favored professional teachers. Assuming validity of criteria, professional teachers identified with the role of Teacher-as-Counselor viewed themselves as more friendly; had more favorable

opinions of pupils, democratic practices, administrators, and other school personnel; had more child-centered educational attitudes; and a more consistent educational philosophy. Professional teachers were rated by trained observers as more responsible-systematic; more skilled in use of teaching media; more competent in non-specific teaching behaviors; and overall more competent. Some non-significant differences favoring provisional teachers indicated more identification with Referrer and Disciplinarian teacher roles, and more "traditional" and "learning-centered" (as opposed to child-centered) viewpoints. A non-parametric sign test indicated that 23 (of a possible 33) differences favoring professional teachers was significant at the .05 level.

Sex differences suggested that women beginning teachers viewed themselves, and were viewed by pupils and observers, as somewhat superior to men. Some significant differences for field of teaching and for place of training, as well as five significant interactions, were obtained.

◆ ◆ ◆

The Effects of a Helping Relationship Experience With Children on Students in Teacher Education

NEWELL T. GILL, ROY KING, RON G. WILBURN, Eastern Kentucky University

A review of the literature suggested a national concern regarding our educational system. Some seem to feel that the breakdown in teacher-pupil communications is at the root of the problem. Solutions point to a more meaningful teacher-education experience such as may be found in a "helping relationship" between a student and a child. The researcher felt such a life experience would have a greater impact in fostering and strengthening attitudes conducive to working with children than an academic experience. The purpose of this study was to explore means for more effectively sensitizing education students to children. It was felt that success in this area would enhance the student's ability to establish good rapport. The hypotheses tested were: 1) The involvement of a student with a child in a helping relationship will enhance the student's general feeling and sensitivity for children. 2) The involvement of a student with a child in a helping relationship will enhance the student's feelings about himself. 3) The involvement of a student with a child in a helping relationship will significantly effect the dogmatism-liberal characteristics of his belief system. 4) Students involved in a helping relationship with Negro children will tend to view themselves and the Negro as being more alike than students not having tutored Negroes. 5) A student's performance on the Minnesota Teacher Attitude Inventory will be significantly affected by the degree to which he is: a) "dogmatic" and b) "secure".

The data for evaluating the hypotheses were collected from a sample of 82 junior and senior education students enrolled in a child study course at Eastern Kentucky University. 29 students worked with one or more white children; 19 with one or more Negro children; and the remaining 34 served as a control. The treatment lasted for approximately eight weeks. The groups were pre- and post-tested on selected personality variables felt to be necessary for the establishment of good rapport. The instruments used were: 1) Maslow's S-I Inventory; 2) Rokeach's Dogmatism Scale; 3) Bogardus Social Distance Scale; 4) A Trait Evaluation Scale; 5) The Minnesota Teacher Attitude Inventory (the MTAI was not available for pretesting).

A complex analysis of variance between groups was the primary statistical method used in analyzing the data. Although the means were most often in the expected direction, only two F's were statistically significant. Females and "liberals" scored significantly higher on the MTAI than males or "traditionalists". The implications of these findings for teacher education were discussed. The failure to find evidence of any significant effect on the experimental groups was attributed to insufficient duration and intensity of the treatment.

3:00 to 3:55 P.M. **SESSION 20.9** **DIVISION C**

CONCEPT LEARNING IV

CHAIRMAN:

Glen Heathers, University of Pittsburgh

PAPERS:

The Effect of Two Types of Information Processing Training in Concept Learning

MARK B. REEVE, DANIEL D. BLAINE and J. L. DUNHAM, The University of Texas at Austin

The major purpose of this study was to investigate the effect of training, under two types of presentation of information processing problems (IPP), upon subsequent performance in a concept learning task (CT). In addition, a paired-associates memory test was constructed and administered to all *Ss* to assess the effect of paired-associates memory upon performance in both the IPP and the CT.

All materials except certain segments of the instructions were programmed for presentation on IBM 1050 computer terminals connected to an IBM 1440 system. The experimental materials were presented in three phases. The first phase (memory test) consisted of items which required *S*s to recall a number (1, 2, 3, or 4) associated with a group of letters from a short list of such associations just previously presented. The second phase (IPP) consisted of the presentation of concept task instances with their appropriate category designation. *Ss'* task was to make a decision about the rejection of a specific hypothesis based upon the information presented in a small number of instances. The same problems were presented to two groups; one received the instances under a simultaneous (SI) condition and the other under a successive (SU) condition. The third phase (CT) consisted of a 4-category conjunctive concept problem with instances presented successively.

Significantly better performance was attained on the IPP in the SI condition; however, memory, as measured by the computer administered test, was not important to *S*'s overall performance in this study. There was no difference in the CT performance between those who were trained in SI and SU conditions. Those who did well on the IPP did significantly better on the CT than those who did poorly in the training phase. This suggests that any enhancement of performance due to similarity of presentation of the SU training and the CT was balanced by superior learning of how to process information in the SI condition.

Information Processing and Concept Learning as a Function of Cognitive Style

WAYNE C. FREDRICK, University of Wisconsin

Witkin and his associates have suggested that cognitive style, specifically the level of field articulation, may be a personality trait which pervades areas of cognitive functioning as well as perceptual activity. Gardner, Jackson, and Messick have stated that selective attention to relevant versus compelling irrelevant stimuli may be the key factor in tests of cognitive style.

The present study was an attempt to verify that field articulation and concept learning are related by the fact that the processing of relevant and irrelevant information is a necessary part of both tasks. To this end, 256 students in grades 6, 8, and 10 were given a test of cognitive style, the Hidden Figures Test (HFT). They were also given a test of information processing, and two concept learning problems.

It was found that performance on the HFT items was significantly correlated with the absolute and relative amounts of relevant and irrelevant information in the items. There seemed to be a developmental trend where younger *S*s were more sensitive to amounts of information in the HFT items than the older *S*s were. Other results of interest were a steady increase in analytical ability with increasing age, and lack of differences in cognitive style between males and females.

*S*s were divided into two groups on the basis of their HFT scores. Those below the median on the HFT (the global *S*s) were significantly poorer in information processing and concept learning than were those above the median (analytical *S*s). An analysis of performance by types of errors showed that the global *S*s made more inclusion, exclusion, and indecision errors than the analytical *S*s. Developmentally, improvement in information processing and concept learning was achieved mainly through a reduction in the number of inclusion errors.

◆ ◆ ◆

The Influence of an Individual's Cognitive Style upon Concept Identification at Varying Levels of Complexity

J. KENT DAVIS, University of Victoria

The purpose of this experiment was to examine the extent to which an individual's cognitive style influenced his performance on concept identification problems of varying levels of complexity. Cognitive style was operationally defined in terms of an individual's performance on the Hidden Figures Test (HFT). It was assumed that subjects (*S*s) able to identify the hidden figures represented an analytic cognitive style, while *S*s unable to identify the hidden figures represented a global cognitive style.

It was hypothesized that individuals experiencing difficulty in locating hidden figures on the HFT would also experience difficulty in identifying concepts in the concept identification task, and that the global *S*s would experience greater difficulty with the more complex concepts.

Three non-overlapping groups of 30 *S*s each were selected from a distribution of HFT scores obtained from 310 senior high school males. Ten *S*s from each of the three levels of cognitive style were randomly assigned to one of three complexity levels (1, 3 or 5 bits of irrelevant information). The concept identification task required *S*s to classify figural patterns, which could vary along as many as seven bi-valued dimensions, into four categories. Each category represented one of four possible combinations of values from two relevant dimensions. Subjects were tested individually and run to a criterion of 16 consecutively correct responses.

Results of the analysis of variance on errors-to criterion indicated significant main effects of cognitive style ($F = 9.51$; $df = 2/72$; p .01), complexity ($F = 18.31$; $df = 2/72$; p .01), and problem type ($F = 20.73$; $df = 1/72$; p .01). Also, two interactions were found to be significant; cognitive style by problem ($F = 4.94$; $df = 2/72$; p. 01) and complexity by problem ($F = 5.82$; $df = 2/72$; p. 01).

As hypothesized, an individual's cognitive style was found to influence his concept identification performance. Individuals identified as analytical on the HFT experienced little difficulty in identifying concepts, while global *S*s experienced considerable difficulty. Individuals falling in the middle of the HFT distribution performed at an intermediate level of performance on the concept identification task. Furthermore, performance was an increasing function of the complexity of the concept identification problems. Although the hypothesized interaction of cognitive style by complexity was not significant, these results suggest that an individual's cognitive style exerts a profound influence on the manner in which he learns.

◆ ◆ ◆

The Effect of Rule Instruction upon the Relationship of Cognitive Abilities to Performance in Multiple-Category Concept-Learning Problems

J. L. DUNHAM and C. VICTOR BUNDERSON, University of Texas at Austin

During the last decade a number of studies have used various factor-analytic procedures to demonstrate the importance of cognitive abilities in various learning tasks. Recent studies have shown that the relative importance of abilities varies differently over trials in concept-learning tasks. An implication of these studies is that there are aspects of the tasks which require the use of particular abilities. If an ability can be linked to a particular aspect of a task, then the manipulation of that aspect should correspondingly affect the relationship of the ability.

The purpose of this study was to ascertain the relationships of relevant abilities to performance in multiple-category concept problems, and to determine if these relationships could be predictably manipulated. Six cognitive abilities—Verbal Comprehension, General Reasoning, Syllogistic Reasoning, Induction, Associative Memory, and Memory Span—were selected on the basis of an analysis of the multiple-category problem. The multiple-category problem can be solved by using simple rules regarding the comparison of two instances. *S* has to form and then use these rules to attain the correct concept. It was hypothesized that instruction and training on these rules would decrease the importance of Induction in these problems. In addition, the abilities that are of more importance after the rules are known should have relationships occurring earlier over a series of problems, than in a group not trained on the rules.

Twelve tests representing the six abilities were administered to 136 high school students. The students were then randomly assigned to two conditions. One group received general instructions. The other received instructions and training on the rules in addition to the general instructions. Both groups were then given a series of nine multiple-category problems. The tests were factor analyzed and vectors representing each problem were projected into the factor space. While some difficulty was encountered in obtaining the expected factor structure, the results in general supported the hypotheses. It was concluded that abilities and task conditions interact, and that the nature of this interaction can be predicted.

TOPICAL INDEX OF PAPER SESSIONS

This brief index is based upon the titles of paper sessions as written by the Division Program Committees. The individual papers have not been indexed by name. References are to Session Numbers.

Topics **Sessions**

- Arts
  - Language 7.7
- Attitudes
  - Toward education 1.7
  - Values and; college students 17.7

- Behavior
  - Administrative, and organizational theory 11.3
  - Classroom, research on 4.4
  - Supervisory, and teacher motivation 4.7
  - Teacher 9.1
  - Teacher, and organizational structure 12.8
  - Verbal 1.8; 5.8; 10.9; 13.5; 14.8; 19.3

- Characteristics
  - Behaviors of school principals and 8.12
  - Group 7.9
  - Of college populations 18.9
  - Of students 20.5
  - Teacher 20.8
- Classroom
  - Behavior, research on 4.4
  - Interaction 1.6, 20.2
- College
  - Populations, characteristics of 18.9
  - Student, adjustment and development 20.6
  - Students: values and attitudes 17.7
- Concept
  - Conservation 5.9, 18.8
  - Learning 4.8; 11.5; 14.5; 20.9
  - Self and pupil performance 6.7
- Conservation
  - Concepts 5.9, 18.8
- Contributed papers 2.6; 8.11; 10.11; 11.7; 14.10
- Counseling
  - Behavioral 6.6
  - Evaluation and micro-procedures 10.8

- Decision Making
  - and planning in educational organization 1.10
- Design
  - Experimental 15.4
- Development
  - Adjustment and, college student 20.6
  - Career, and vocational interests 19.6
  - Curriculum 18.7
  - Measuring instruments, of educational 7.4
  - Psycho-social 12.6
  - Pupil, growth relationships and 16.6
- Disadvantaged
  - Students, curricula for 13.2
  - Students, language instruction for 12.5
  - Teaching the 6.2, 17.4

- Education
  - Attitudes toward 1.7
  - Compensatory and vocational 7.8
  - Early childhood 3.5
- Educational Outcomes
  - Variables related to 17.6
- Empirical Test Research 19.5
- Environments
  - Learning, studies of 20.3
- Evaluation
  - Curriculum 20.4
  - Of teaching strategies 10.5, 18.6
- Experimental Design 15.4

- Individual Differences
  - Dimensions of 10.10
  - Learning and 2.5, 13.6
- Instruction
  - Language, for disadvantaged students 12.5
  - Mathematics 8.3
  - Programmed 6.4, 8.6, 14.6
- Integration
  - School 20.1
- Intellectual Processes
  - Measurement of complex 13.4
- Interaction
  - Classroom 1.6, 20.2
- Interests
  - Vocational, and career development 19.6

Topics **Sessions**

Language
  Arts 7.7
  Instruction for disadvantaged students 12.5
Learning
  Concept 4.8; 11.5; 14.5; 20.9
  Environment, studies of 20.3
  From prose 3.4, 11.1
  Individual differences and 2.5, 13.6

Mathematics
  Instruction 8.3
  Science and 3.6
Measurement
  Of complex intellectual processes 13.4
Measuring Instruments
  Development of educational 7.4
Memory 9.3
Micro-Teaching 7.2
Models
  Research, Computer-Related 1.1
Motivation
  Teacher, and supervisory behavior 4.7

Non-American Subjects
  Studies of 20.7

Papers
  Contributed 2.6; 8.11; 10.11; 11.7; 14.10
Performance
  Pupil, and self-concept 6.7
Planning
  And decision making 1.10
Policy Determination
  Educational, political aspects of 1.11
Prediction
  Academic Success 2.2
Problems
  Special, in administrative research 4.6
Problem Solving
  Thinking and 15.1
Programmed Instruction 6.4, 8.6, 14.6
Prose
  Learning from 3.4, 11.1

Topics **Sessions**

Reading
  Beginning 5.3
  Research 7.6
Reinforcement 9.2, 15.2

School
  Board-superintendent relationship 16.3
  Integration 20.1
  Principals, characteristics and behavior of 8.12
Science
  Mathematics and 3.6
Social Studies 3.7
Socioeconomic Status 7.5, 12.7
Special Groups
  Research on 6.5
Statistical
  Techniques 13.3
Success
  Predicting academic 2.2

Teacher
  Behavior 9.1
  Behavior and organizational structure 12.8
  Characteristics 20.8
  Prospective, studies of 19.4
Teaching
  Micro- 7.2
  Research on 15.3
  Strategies, evaluation of 10.5, 18.6
  The disadvantaged 6.2, 17.4
Techniques
  For interpretation of results 14.1
  Statistical 13.3
Testing
  Research on 7.3, 11.4, 14.9
Theory
  Organizational and administrative behavior 11.3
  Test 16.4
Thinking
  Problem solving and 15.1

Verbal
  Behavior 1.8; 5.8; 10.9; 13.5; 14.8; 19.3

INDEX OF PARTICIPANTS

The following index concerns the authors, co-authors, sponsors, and chairmen of the paper sessions of the 1968 A.E.R.A. Annual Meeting. Reference is to Session Numbers.

Names Sessions

Aaron, Robert L. 12.7
Agan, Raymond J. 7.8
Alexakos, Constantinos E. 11.4, 13.3
Alexander, William M. 18.7
Alkin, Marvin C. 17.6
Allen, Dwight W. 9.1
Allender, Jerome S. 10.5
Allman, Thomas S. 10.10
Almy, Millie 3.6
Amarel, Marianne 20.1
Amidon, Edmund J. 20.2
Anderson, Dennis L. 5.8
Anderson, Harry E. Jr. 10.10; 15.4; 20.3; 20.7
Anderson, James G. 13.2
Anderson, Richard C. 2.6; 9.1; 9.2; 10.9; 20.2
Andre, Thomas 7.9
Arnold, Richard D. 6.5, 12.5
Ashcraft, Rock 20.4
Ausubel, David P. 3.4
Avital, S. M. 8.3

Bahlke, Susan J. 20.1
Baker, Eva L. 15.3
Baker, Frank B. 1.1, 15.4
Balow, Irving H. 3.5
Bargar, Robert R. 10.11
Barrett, Thomas C. 3.4
Barritt, Loren S. 2.6, 12.7
Bashaw, W. L. 1.7, 20.3, 20.7
Bauer, Roger 2.2
Beaber, James D. 6.6, 8.11
Beaird, James H. 16.6, 18.7
Beauchamp, George A. 3.5
Beittel, Kenneth R. 2.6
Benning, James J. 12.6
Benson, Charles S. 17.5
Bentley, Joseph C. 18.9
Bentley, Ralph R. 12.3
Bentzen, Mary M. 11.3
Berger, Robert J. 3.5
Berlak, Harold 1.6
Besag, Frank P. 7.8
Berzonsky, Michael D. 4.8
Bhushan, Vidya 8.3

Names Sessions

Bickley, A. C. 14.9
Bishop, Robert 14.10
Blaine, Daniel D. 14.5, 20.9
Bledsoe, Joseph C. 20.8
Blommers, Paul J. 19.4
Blood, Ronald E. 8.12
Blumberg, Arthur 4.7
Bodenstab, Ernest W. 19.4
Boersma, Frederick J. 2.5, 7.3
Bott, William K. 19.4
Bower, James R. 20.7
Bower, W. Scott 13.4
Bowers, Norman D. 13.6, 20.4
Bracht, Glenn H. 7.3
Bradford, C. H. 13.3
Braskamp, Larry A. 18.9
Braun, John R. 17.7
Brega, Evelyn 20.4
Bridges, Edwin M. 8.12
Briner, Conrad 1.11
Brinkmann, Erwin H. 20.3
Brookover, Wilbur B. 20.1
Brown, Bob Burton 4.4; 11.7; 15.3; 19.4
Brown, Bobby R. 8.6
Brown, Roscoe C. 7.8, 10.5
Brumbaugh, Robert B. 11.3
Bunderson, C. Victor 3.6, 20.9
Burmeister, Lou E. 10.9
Busby, Walter A. 14.8
Busse, Thomas V. 15.1
Byers, Joe L. 1.8, 4.8, 5.8

Cahen, Leonard S. 7.4, 20.4
Caldwell, Edward 19.3
Carlson, Jerry S. 5.9
Carlson, Richard 11.3
Carnie, George M. 16.3
Carroll, J. Douglas 13.5
Carry, L. Ray 14.1
Carton, Aaron S. 11.1
Chapin, June R. 19.4
Chase, Clinton I. 15.4
Clarizio, Harvey 12.6
Clark, Philip M. 13.4
Clark, Richard M. 19.6
Clawar, Harry J. 14.1

Names **Sessions**

Cleary, T. Anne 15.4
Cohen, Sandra R. 20.1
Cole, Henry P. 18.8
Conley, James 20.8
Conway, James A. 20.4
Cook, Harold 15.2
Cooper, Carin S. 1.8
Costin, Frank 20.5
Cotton, John W. 4.8
Coughlan, Robert J. 4.7
Cowles, Milly 7.7, 13.2, 17.4
Cox, Richard C. 11.4, 16.4
Craig, Robert C. 15.1
Crawford, William R. 7.4
Crispin, David B. 8.11
Cronin, Joseph M. 4.6
Crowley, Robert J. 1.7

Daniel Kathryn B. 9.2, 13.2
Davidson, Robert E. 10.9
Davis, Frederick B. 7.6
Davis, Gale 3.5
Davis, J. Kent 20.9
Davis, O. L. Jr. 7.7, 20.8
Day, David E. 12.5
Day, H. I. 13.4
D'Costa, Ayres 19.6
Deno, Stanley L. 14.5
DeSousa, Albert M. 7.7
Deterline, William A. 6.4
Dettre, John R. 20.4
Devor, Geraldine 9.2
DiLorenzo, Louis T. 17.4
Dixon, W. Robert 4.4
D'Oyley, Vincent R. 7.4
Dooley, B. J. 3.7
Doppelt, Jerome E. 11.4
Duell, Orpha K. 19.3
Duncan, James K. 9.1
Duncanson, D. L. 17.6
Dunham, J. L. 3.6, 4.8, 14.5, 20.9
Durnin, John H. 14.5
Dutter, Donald R. 10.5
Dyer, Calvin O. 6.5
Dyer, Jean L. 13.4

Eash, Maurice J. 18.7
Ebel, Robert L. 15.4, 16.4
Edwards, Joseph C. 11.5
Egermeier, John C. 6.2
Egner, Joan R. 8.12
Ehri, Linnea 12.7

Names **Sessions**

Eidell, Terry L. 4.7
Eisner, Elliot W. 12.5
Emans, Robert L. 6.2, 12.5
Emmer, Edmund T. 1.6
Emrick, J. A. 11.5
Endo, George 10.5
Ericksen, Gerald L. 3.6
Erickson, Edsel L. 20.1
Evans, D. R. 13.4
Evans, Keith 10.5

Faust, Gerald W. 3.4, 6.4
Fay, Peter P. 1.7
Fears, Emory B. 7.9
Feldhusen, John F. 2.2, 12.6
Ferreira, Joseph 1.10
Feshbach, Norma D. 9.2, 15.2, 17.4
Feshbach, Seymour 17.4
Findley, Warren G. 7.4, 20.3
Fisher, Gerald A. 4.4
Flatter, Charles H. 9.2
Fleming, James T. 7.5
Fleming, Malcolm 5.8
Flizak, Christopher W. 12.8
Follman, John C. 13.4
Foulke, Emerson 19.3
Frankiewicz, R. G. 13.5
Frase, Lawrence T. 3.4, 11.1, 18.6, 19.3
Frederick, Franz J. 10.8
Frederick, Wayne C. 20.9
Fredericks, Harold D. 18.7
Fredericks, Patricia 8.6
Frederiksen, Norman 19.5
Freshley, Dwight L. 10.10
Friedman, S. Thomas 20.5
Frost, Joe L. 13.2
Fry, Edward 7.7
Furukawa, James 6.4

Gage, Gerald E. 18.7
Gaite, A. J. H. 3.4
Garber, Herbert 10.11
Gardner, Eric F. 19.3
Garfunkel, Frank 6.2, 14.10
Garverick, Charles M. 9.2
Georgeoff, P. John 20.1
Gibson, R. Oliver 16.3
Gilbert, James E. 6.4
Gill, M. P. 7.4
Gill, Newell T. 20.8
Gillooly, William B. 6.4
Girault, Emily S. 3.7

Names	Sessions
Gladney, Mildred R.	12.5
Glass, Gene V.	4.4
Glock, Marvin D.	14.5
Goldberg, Isadore	10.8
Goodkind, Thomas B.	7.2
Goodwillie, Richard O.	20.5
Goodwin, Harold I.	1.11
Goodwin, William L.	10.5
Gordon, Jack H.	18.6
Greene, Frank P.	7.6
Gregor, A. J.	6.5
Grinder, Robert E.	5.9
Grobman, Hulda	8.3
Grossberg, John M.	6.6
Guertin, Wilson H.	20.7
Gustafson, H. W.	14.6
Gustafson, Robert	17.6
Haddox, Victor G.	10.8
Halamandaris, Pandelis G.	5.8
Hall, James W.	10.9
Hall, Vernon C.	19.3
Haller, Emil J.	4.6
Hamachek, Don E.	20.8
Hambleton, Ronald K.	11.4, 14.9
Hampton, Claudia H.	1.7
Hanbury, Elizabeth B.	7.7
Hansen, Duncan N.	2.5
Harris, Albert J.	14.9, 17.4
Harrison, Forrest W.	6.7
Harrison, Grant V.	18.6
Hawkes, Thomas H.	7.9
Hawkins, Richard	10.8
Heath, Robert W.	7.4
Heathers, Glen	20.9
Heckman, Robert W.	13.6
Hedley, Carolyn Neal	17.7
Heller, Robert W.	12.8
Helm, Carl E.	14.9
Hemberger, Lance W.	20.3
Henderson, Edmund H.	16.6
Hickcox, Edward S.	16.3
Hill, Russell A.	12.8, 20.2
Hiller, Jack H.	4.4
Hillery, Milton C.	3.5
Hines, Vynce A.	18.7
Hively, Wells II	20.3
Ho, Wai-Ching	10.10
Hogben, Michael	10.5
Hollen, T. T. Jr.	3.6
Holmes, James N.	1.10
Hopkins, Kenneth D.	7.3

Names	Sessions
Hornbostel, Victor O.	6.2
Horowitz, Bernard	20.5
Hosford, Ray E.	6.6
Houser, Ronald L.	15.3
Houston, Camille	7.6
Houston, Tom	1.1
Howe, Michael J. A.	9.3
Hubbard, Duke B.	1.11
Huitema, Bradley	18.7
Hummel, Raymond C.	7.8
Hunkins, Francis P.	3.4
Hunter, Elizabeth	14.8
Hurd, Donald E.	14.8
Husek, T. R.	14.10, 19.5
Hyman, Ronald	20.1
Ilika, Joseph	3.6
Ivany, J. W. George	3.6
Jackson, Philip W.	9.1
Jacobson, Milton D.	1.1; 6.6; 7.7; 8.11; 10.8; 14.6; 16.4; 18.6
Jeffryes, James	8.3
Jenkins, David M.	1.7
Jenkins, Joseph R.	19.3
Jennings, William N.	7.3
Joe, George W.	15.4
Johnson, Charles E.	20.3
Johnson, David W.	17.7
Johnson, Homer M.	16.3
Johnson, Joseph C. II	7.7
Johnson, M. Clemens	1.1
Johnson, Paul E.	13.5, 18.8
Johnson, Ronald E.	9.3
Johnson, Thomas J.	14.10
Judy, Chester J.	2.2
Juola, Arvo E.	16.4
Jurowski, Edward S.	10.5
Kallenbach, W. Warren	7.2
Kameya, Lawrence I.	2.5
Karis, Charles	6.4
Karraker, Ronnie J.	15.2
Keislar, Evan R.	3.5, 11.5, 12.7
Kennedy, John J.	6.6
Kenney, James B.	1.7; 8.11; 16.3; 17.6
Kent, Alvin	6.4
Kim, Yungho	20.7
King, Roy	20.8
Kingston, Albert J.	13.5
Klausmeier, Herbert J.	4.8
Klemmack, David L.	6.7

Names	Sessions
Kliebard, Herbert M.	13.2
Knoell, Dorothy M.	13.4
Koenke, Karl R.	11.1
Koff, Robert H.	7.9
Kohl, Donald	18.7
Krahmer, Edward F.	20.6
Krasno, Richard M.	9.1
Lambie, Dolores Z.	6.2
Lander, Dorothy E.	7.8
Landrum, William	1.7
Lankton, Robert	14.1
Lantz, Donald L.	19.4
Lawrence, Clifford J.	16.3
Leaverton, Lloyd	12.5
Lennon, Roger T.	7.3
Levin, Joel R.	1.8
Lezotte, Lawrence W.	1.8
Lieberman, J. Nina	20.3
Lillich, Joseph M.	6.2
Linden, Kathryn W.	15.3
Lingren, Ronald H.	3.5
Linn, Robert L.	15.4
Lippert, Henry T.	13.6, 14.5
Long, Barbara H.	16.6
Lorentz, Jeffrey	17.6
Loykovich, Joan M.	19.5
Lupton, Daniel E.	8.11
Lynch, William W.	14.5
MacGinitie, Walter H.	13.5
Magoon, Jon	16.4
Maguire, Thomas O.	4.4
Mahan, David J.	16.3
Mahan, J. Maurice	20.4
Mahoney, Leo G.	1.10
Mann, Philip H.	6.6, 8.11
Manning, Winton H.	1.7
Mansergh, Gerald G.	12.6, 20.5
Marquit, Lawrence J.	4.7
Marsh, Luther A.	8.6
Marshall, Daniel W.	9.3
Massad, Carolyn Emrick	7.5
Mathis, B. Claude	18.8
Mattleman, Marcienne S.	12.5
Mattson Dale E.	7.9
McCarthy, Edward H.	2.6
McClave, James T.	19.4
McFadden, Jack D.	18.9
McGee, Robert G.	8.3
McLaughlin, Kenneth F.	14.9
McMillan, Joseph H.	20.1
McNaughton, Anthony H.	7.4
McNeil, Keith A.	19.5
McQuitty, Louis L.	13.3
Medley, Donald L.	20.2
Mehrens, William A.	2.2, 11.7
Melaragno, Ralph J.	17.4, 18.6
Meredith, Gerald M.	18.9
Mermelstein, Egon	5.9
Merrifield, Philip R.	19.5, 20.6
Merrill, M. David	9.3
Merritt, James W.	3.7
Metzner, Seymour	20.6
Meux, Milton O.	10.5
Meyer, Edwina	5.9
Michael, Calvin B.	10.11
Mickelson, John M.	10.11
Monson, Dianne L.	7.9
Moore, J. William	20.4
Morrison, Coleman	17.4
Morrow, Richard G.	17.6
Morse, P. Kenneth	15.3
Moskowitz, Gertrude	20.2
Mulcahy, Gloria A.	9.2
Murray, C. Kenneth	9.1
Murray, Frank B.	18.8
Nagel, Elwyn H.	6.7
Nagel, Thomas S.	14.6
Nally, Thomas P.	16.6
Nash, Shirlyn	18.7
Neale, Daniel C.	17.7
Nelsen, Edward A.	12.6
Nelson, Ronald	1.7
Nelson, Suzanne	3.5
Newell, John M.	20.4
Newmark, Gerald	17.4
North, Robert D.	11.7
Novotney, Jerrold M.	11.3
Ober, Richard L.	4.4
O'Brien, Thomas C.	15.1
OBryan, Kenneth G.	2.5
Ohnmacht, Fred W.	7.6
Okorodudu, Corahann	10.11
Olson, Arthur V.	7.6
Olson, Gary F.	8.12
Olson, Norinne H.	7.6
O'Neil, Harold F.	2.5
Oppenheim, Don B.	10.10
Orme, Michael E. J.	15.2
Osborne, R. Travis	6.5, 20.3
Otto, Wayne	1.8, 3.4, 7.6, 14.8

Names	Sessions
Overstreet, J. Douglas	4.8
Owoc, Paul	14.10
Page, Ellis B.	1.1
Panos, Robert J.	6.7, 14.1
Parsons, Patricia J.	19.3
Patterson, Wade	11.3
Patton, Don C.	1.11
Patton, Robert W. Jr.	1.7
Paulus, Dieter H.	14.9
Penfield, Douglas A.	16.4
Perlberg, Arye	7.8
Peyser, Turkan K.	20.7
Pezzoli, Jean A.	11.1
Phillips, Beeman N.	6.5, 10.10
Pierce-Jones, John	8.6
Planisek, Robert W.	20.6
Plattor, Emma E.	13.2
Plattor, Stanton D.	13.2
Ploski, Harry	7.8
Popham, W. James	15.3
Porter, Andrew C.	15.4
Porter, Douglas	15.2
Portuges, Stephen H.	15.2
Pounds, Haskin B.	16.3
Powell, Evan R.	20.2
Prentice, Joan L.	1.8, 5.8
Prichard, Allyn	20.7
Pruzek, Robert M.	11.7
Purkey, William W.	10.8
Purnell, Richard F.	15.2
Purrington, Gordon S.	11.3
Pyatte, Jeff A.	8.6
Pyle, Thomas W.	14.5
Ram, Yael	7.8
Ramos, Robert	7.3
Ramsay, James G.	4.8
Rarick, G. Lawrence	14.8
Raven, Ronald	18.8
Rayder, Nicholas F.	17.7
Reed, James C.	5.3
Reeve, Mark B.	20.9
Remstad, Robert C.	3.5
Rentz, Robert R.	1.7, 7.9 ,17.6
Renzulli, Joseph S.	14.9
Resnick, Lauren B.	5.8, 12.7
Reynolds, Richard J.	9.1
Richards, James M. Jr.	18.9
Rippey, Robert M.	4.6
Ripple, Richard E.	2.5
Ritts, C. Allen	20.4

Names	Sessions
Roberts, Art	4.4
Roberts, S. O.	10.10
Roberts, Tommy L.	10.8
Robinson, Helen M.	5.3
Rogers, Bruce G.	11.7
Rohwer, William D. Jr.	1.8, 12.7
Romberg, Thomas	11.4
Rose, Gale W.	8.12
Rosemier, Robert A.	11.5
Rosenbaum, Edward	12.6
Ross, George H.	1.10
Rossman, Jack E.	18.9
Rothkopf, Ernst Z.	13.5, 14.6
Rozran, Andrea R.	11.1
Rubincam, Irvin R.	10.5
Rusch, Reuben R.	19.6
Ryan, Frank L.	8.3
Ryan, James J.	3.6, 6.7, 7.3
Ryan, T. Antoinette	6.6, 10.8, 16.6, 17.7
Safar, Dwight	13.2
Salomon, Gabriel	7.5
Salter, Ruth	17.4
Samuels, Joanna	4.7
Samuels, S. Jay	4.7, 10.9
Sandborgh, Raymond E.	13.5
Sands, Harold R.	8.6
Sassenrath, Julius M.	9.3, 14.8
Sawin, Enoch I.	7.4
Scandura, Joseph M.	8.3, 14.5
Schlesser, George E.	1.7
Schoenfeldt, Lyle F.	2.2
Schuerger, J. M.	13.5
Schumer, Harry	7.8, 14.6, 18.9
Schutz, Samuel R.	11.5, 12.7
Scott, Owen	3.7
Seibert, Warren F.	11.7, 13.6
Semmel, Melvyn I.	9.3
Sergiovanni, Thomas J.	4.6
Shapiro, Bernard J.	15.1
Shaver, James P.	7.2, 13.6
Shaw, Robert A.	1.1, 14.6
Shaycoft, Marion F.	13.3
Shepler, Jack	11.4
Sherman, Vivian S.	19.6
Sherrill, David M.	20.5
Sheverbush, Robert L.	19.5
Shoemaker, Harry A.	14.6
Shulman, Lee S.	6.7
Sieber, Joan E.	2.5, 7.5
Simpson, Gwenn	19.3
Singh, B.	14.9

Names	Sessions
Sirotnik, Ken	14.10, 19.5
Sister Josephina	18.6
Sjogren, Douglas D.	15.1, 19.6
Skipper, Charles E.	20.8
Slotnick, Henry B.	15.1
Smith, Charles E.	11.4
Smith, Gerald R.	1.10
Smith, Helen K.	5.3
Smith, Louis M.	20.2
Smith, Martin E.	11.7
Smith, Richard	14.8
Snow, Richard E.	13.6
Soar, Robert S.	1.6
Sokol, John	1.10
Sorenson, Garth	10.8
Sowards, G. Wesley	3.7
Spady, William G.	20.6
Spaights, Ernest	6.6
Spielberger, Charles D.	2.5
Stager, Mary	3.4
Stakenas, Robert G.	14.10, 20.6
Stanfield, Robert E.	18.9
Stanley, Julian C.	13.3
Starks, David D.	2.2
Steg, Doreen E.	9.2
Stern, Harris W.	13.2
Stoffel, William N.	4.4
Sugerman, D.	8.3
Suzuki, Nancy	1.8, 12.7
Taft, Jerome	7.4
Talmage, Harriet	18.7
Tanaka, Masako N.	13.6
Taranto, Maria	1.7
Tarte, Robert D.	2.6
Tatham, Susan M.	11.1
Thomas, Michael L.	1.11
Thomas, Sally A.	11.5
Thomas, Walter L.	17.7
Thompson, Glen R.	13.6
Thorn, Elizabeth A.	5.3
Thurner, Ronald	13.5
Thurston, John R.	12.6
Tiemann, Philip W.	6.4, 8.6
Tillman, Murray H.	6.5, 14.10
Tobias, Sigmund	2.5
Tobiason, Ray	12.8
Tondow, Murray	12.6
Toothaker, Larry E.	15.4
Towler, John O.	5.9
Traub, Ross E.	14.9
Travers, Kenneth J.	7.4

Names	Sessions
Traylor, Elwood B.	9.1
Trueblood, Cecil R.	18.8
Tuckman, Bruce W.	1.6, 2.6
Twyman, J. Paschal	6.2
Unks, Nancy J.	11.4
Urbas, Raymond	6.2
Van Mondfrans, Adrian P.	9.2
Verplanck, William S.	7.3
Vickery, Tom Rusk	11.7
Vinsonhaler, John F.	2.2
Vogel, Francis X.	20.4
Voth, John A.	7.2
Wahlstrom, Merlin W.	7.3
Walberg, Herbert J.	19.6
Wallace, Gaylen R.	6.2
Wallen, Norman E.	7.4
Wallin, Herman A.	4.6
Ward, Robert	20.5
Watkins, Foster	1.7
Weaver, Wendell W.	13.5, 14.9
Webb, Jeaninne N.	15.3
Weber, Wilford A.	1.6
Weikart, David P.	6.2
Weintraub, Samuel	5.3
West, Leonard J.	2.6
Wheatley, Grayson H.	18.8
White, William F.	7.9, 10.10, 12.7, 13.5
Wilburn, Ron G.	20.8
Wiley, Glenn E.	1.11
Williams, Charlotte	6.5, 14.10
Williams, Joanna P.	5.8, 10.9
Williams, Reed G.	14.5
Williamson, John A.	15.3
Williamson, John H.	2.5
Willingham, Warren W.	2.2
Wilson, Helen K.	6.6
Wilson, James W.	8.3, 14.1
Winefordner, David W.	19.6
Wittrock, M. C.	11.5
Wodtke, Kenneth H.	8.6
Wolf, Richard M.	16.6
Wolf, Willavene	7.5
Wolff, Joseph L.	11.5
Wolfson, Bernice J.	18.7
Wood, Sam	4.4
Wrightsman, Lawrence S.	1.7
Yamamoto, Kaoru	20.8

Names	Sessions	Names	Sessions
Yarborough, Betty	7.7	Young, David B.	7.2
Yeatts, Pearline Peters	16.6	Young, John J.	16.4
Yee, Albert H.	7.5		
Yonge, George D.	9.3	Zussman, Hedy	10.5